30c

Hazen Snarr
164 "C" Street
Salt Lake City

Territory

Original terr — 892,135 sq
Paris Treaty 1783

Louisiana Purchase 827,987
1803

Florida — 71101
Sp $500000000
1819

Oregon 286541

~~1847~~

Dis 1792
Ex 1802

Settlement 1811
Spain 1813
Russia 1825

$$\begin{array}{r} 55000 \\ 45 \\ \hline 275000 \\ 220000 \\ \hline 2475,000 \end{array}$$

Q. What are the two controlling influences the government exerts over production

M 1, Costs

Fixed & overhead costs { Taxes
Insurance
Depreciation

Operating Ex { lights
coal
heat
etc.

Particular { Increase or decrease
of production costs
determined by increase
or decrease of business.

Texas
1845— 389,166

1850 Tex to U.S. all claims
to N.M for 10000000

. What relation between
ur resources and sources of
ror emigration. ?
. What baring has this on
class conciousness
3 What has this to do with
distribution of wealth in U. S.
= ~~what relation~~ With
passing of free land how
have our land problems
changed.

II Others.

+ Mexican cession 529,189
pop conquest + Purchase
nd $15 000,000 1848
M. C.
Gadsden 29607
1853. 10,000,000

Alaska Pur. 590 884
From Russia 1867 $7200000
Hawaii 6,449
annexed 1898

ppa.	100,000	200000
F	100,000	100,000
P	500,000	1,000,000
~~C.~~	~~200,000~~	

~~100,000~~	~~200,000~~
100,000	100,000
500 000	1000000
600000.	1200000
1,~~200000~~	23~~00000~~

Spanish cessions

Porto Rico	3,435
Phillipians	115 026
Guam	210

By Paris Treaty 1898

$ 20,000,000

Samoa
Tutuila } 77.
Convention GB.
1899

Q.1. Are benefits of nat. boundries sufficient to off set societies econ. losses because of them?

2. Why is spirit of nationalism as strong as it is today? With what econ results?

III ~~What will cause us to~~ Which is more important for immediate future; increase-production or improved distribution of wealth? Why

Ans in writing

~~what will cause us to conserve our resources. Is the prospect a happy one.~~

~~Write to hand in.~~

Panama 527

Convention Rep of Panama

$10000000 1904

ELEMENTARY ECONOMICS

annuity of $250000

~~St Thos~~

St J

St Croix

Virgin Is. 132

from Denmark

taxing government.

25,000,000.05 1916

U.S. Area. 3,088,519

Outlying 716,740

3,805,259

Bland in Egypt — Madrews

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN & CO., LIMITED
LONDON · BOMBAY · CALCUTTA
MELBOURNE

THE MACMILLAN CO. OF CANADA, LTD.
TORONTO

ELEMENTARY ECONOMICS

BY

FRED ROGERS FAIRCHILD
PROFESSOR OF POLITICAL ECONOMY

EDGAR STEVENSON FURNISS
PELATIAH PERIT PROFESSOR OF POLITICAL AND SOCIAL SCIENCE

NORMAN SYDNEY BUCK
ASSISTANT PROFESSOR OF POLITICAL ECONOMY

YALE UNIVERSITY

VOLUME I

New York
THE MACMILLAN COMPANY
1926

All rights reserved

PRINTED IN THE UNITED STATES OF AMERICA

Copyright, 1924 and 1925,
By FRED ROGERS FAIRCHILD.

Copyright, 1926,
By THE MACMILLAN COMPANY.

Set up and electrotyped. Published May, 1926.

Norwood Press
J. S. Cushing Co. — Berwick & Smith Co.
Norwood, Mass., U.S.A.

Bus/Ec — Bus/Ec, PK

PREFACE

A newcomer amidst the present goodly company of general texts in economics may fairly be called upon for an introduction justifying its appearance. This book is the result of a good many years' experience in teaching economics to beginners in Yale University. That experience has shown us the merits and the defects of the many books on economic principles which are available for use as texts in college classes. It has convinced us of the desirability and likewise of the possibility of preparing a more useful text for the purpose of helping college students to a knowledge of the fundamental facts and principles of the economic world in which they live. That is the sole purpose of this book. Our contribution to economic theory, if anything at all, is slight. The book is not written for the edification of the mature student of economics. It is strictly a book for beginners and the teachers of beginners.

Adherence to this purpose has indicated certain rules for our guidance and imposed certain restrictions upon us. We have sought generally to avoid controversial subjects. As a rule the treatment has been confined to topics upon which the science of economics may fairly be said to have reached definite conclusions. Occasionally, when this rule could not be followed, we have stated, somewhat dogmatically perhaps, what we consider sound conclusions as to subjects upon which there may still be real difference of opinion among economists of authority. We do not believe that the beginning student profits greatly from that type of discussion which presents arguments on both sides of a controversial question and leaves him to draw his own conclusions. Such matters belong in the more advanced courses. There is plenty of material to occupy fully the elementary course in economics without going into those fields upon which economic science has not reached at least fairly definite conclusions.

This limitation of the field makes easier the strictly scientific treatment to which we have sought to adhere. Our purpose is to present the facts and the principles of economics — to disclose the truth. With the ethical or moral problems which arise upon this economic foundation we are not concerned, except only as prevailing judgments upon such matters may themselves become factors in economic activity. Our purpose is to aid the student to discover how things are, not how they ought to be. We have no case to establish, and we offer no propaganda.

Addressed as it is to beginning students, this is essentially an elementary book. We have sought to spare the student the confusion that comes from taking for granted knowledge which he cannot fairly be assumed to possess, and the treatment is correspondingly elementary. We have sought, by giving care to such things as definition of technical terms, sequence of topics, and simplicity of style, to enable the beginning student to follow the discussion without undue difficulty. It is not to be denied that certain parts of the subject of economics are difficult. We have made no attempt to escape such difficulties, either by avoiding the difficult topic or glossing it over with a superficial discussion. Rather have we relied upon simple and elementary treatment to remove unnecessary causes of confusion, in the belief that so presented these topics will not prove beyond the capacity of the student.

The plan of the book has dictated the continuous combination of theoretical analysis with historical narrative and discussion of practical problems. It is our conviction that this is a sound pedagogical rule for an elementary treatise. Theory must be illustrated and justified by showing its relation to practical affairs; facts must be made significant and interesting by showing their conformity to general principles. "Theory" and "practice" should, we believe, go always hand in hand.

It is today quite generally recognized that the average college student does not possess that knowledge of the economic environment which is necessary as a foundation for the study of economic principles and problems. A section (Part I) devoted to the study of the modern economic organization and its development, stressing the coöperative nature of modern society and the importance of price as a motivating force, is therefore made an integral part of this book.

The problems which will be found at various points do not conform to any rigid rule. They have been introduced only at those points where it was felt that they would contribute substantial aid to the student in his effort to grasp the principles expounded in the preceding text or would give useful training in economic reasoning. These problems are obviously illustrative rather than comprehensive, and the teacher can readily modify or expand them or ask the student himself to construct similar ones.

At the end of each of the major subdivisions of the book will be found a classified list of books which are suggested for the student's further reading. These lists are, of course, not comprehensive, and they do not embrace works of a highly advanced or specialized or technical nature. They are intended merely to present for the choice of the student certain books which he will find interesting and profitable reading in connection with his first study of economics.

As compared with the majority of textbooks in economics, this book will appear quite long. It has been written with the idea of comprising the bulk at least of the reading to be assigned as the required work in a college course involving three exercises a week for a full year. We are of the belief that the present-day importance of economics justifies the devotion of that much time to the general elementary course.

These in brief are the more important principles which have guided us in harmony with our aim to give to those students who may not pursue the study of economics further a broad knowledge and understanding of the economic world of today and to others a firm foundation for their further study of economics and allied subjects. It is in such matters as are here noted that our contribution, if any, is to be found.

While we have had in mind exclusively the needs of college students, we are nevertheless bold enough to believe that the requirements thus imposed upon us will have made this book useful also to those more mature readers who, not having had the advantage of college instruction in economics or feeling that their present notions are perhaps vague and antiquated, may wish to read a simple elementary exposition of the fundamental facts and principles of economics.

We make grateful acknowledgment of the invaluable assistance which we have derived from our colleagues in the Economics Department of Yale University. Though the actual task of authorship has fallen to the three undersigned, this book represents the combined contributions, direct and indirect, of a large group of teachers who for many years have been closely associated in the task of teaching elementary economics to the students of Yale University. To all of these, past as well as present colleagues, we acknowledge our debt. In particular we are pleased to record our indebtedness to Professor Irving Fisher, whose *Elementary Principles of Economics* has long served as one of the principal textbooks in the elementary economics course in Yale College and whose influence upon the authors, his former students, has been far-reaching; likewise to Professor Clive Day, who has given valuable suggestions in the field of economic history, to Professor Winthrop M. Daniels for his assistance in the preparation of the chapters on railroads, to Professor Ray B. Westerfield for contribution of valuable material on the subject of risk and insurance, and to Professor Maurice Davie, who wrote the chapter on immigration and aided us by assembling the materials upon population.

Our obligation is by no means confined to our Yale colleagues. First of all we must acknowledge the influence upon our work exerted by Professor F. W. Taussig, through his *Principles of Economics*, which has long served as the principal text in the elementary course in Yale College. For extremely helpful criticisms and suggestions, particularly with reference to the treatment of price, we are indebted to Professors Herbert J. Davenport, Wesley

C. Mitchell, Harry G. Brown, and A. B. Fairchild. General acknowledgment is finally due the authors of the standard works on many subjects embraced within the broad field of general economics, upon whose writings we have leaned heavily and far beyond the possibility of specific acknowledgment.

FRED ROGERS FAIRCHILD
EDGAR STEVENSON FURNISS
NORMAN SYDNEY BUCK

NEW HAVEN, CONNECTICUT
March, 1926

CONTENTS

VOLUME I

PART I. THE ECONOMIC ORGANIZATION

PAGE

CHAPTER I. FOUNDATIONS OF THE SCIENCE OF ECONOMICS . . 3

Human wants. Food. Clothing. Shelter. Other wants. Intermediate goods. No limit to the wants of man. Nature, the provider. Economics. Economics as a science. Utility. Definition of wealth: utility. Materiality. Wealth must be owned. Are human beings wealth? How to identify wealth. The problem of definition The requirements of a scientific definition. Income. Costs. Net income. Property. Division of property. Property in free persons. Wealth and property. Documents in evidence of property. Wealth and income: fund and flow. Value as a common denominator. Transfer and exchange. Barter and money exchange. Value. Price. Value and price. The unit of value.

CHAPTER II. THE FACTORS OF PRODUCTION 26

Production. The factors of production. Land. Labor. Race and inheritance. Health and energy. Training and education. Capital. Capital in production. The accumulation of capital. Saving. The form of capital instruments. Invention. The capitalist class. The entrepreneur.

CHAPTER III. INDUSTRIAL STAGES 42

The household system. The manor. The Handicraft Stage. Merchant gilds. Craft gilds. Municipal regulation of commerce. The beginning of specialization. Growth of freedom. Judgment of the town economy. The domestic system. The factory system.

CHAPTER IV. THE INDUSTRIAL REVOLUTION 55

Why the Industrial Revolution began in England. Coal and iron. The textiles. Spinning. Weaving. The use of power and the growth of factories. The reluctance to establish factories. Developments in other lines. Transportation. Summary of the effects of the Industrial Revolution. The effect of the Industrial Revolution on the laborers. The Industrial Revolution in America.

PAGE

CHAPTER V. DIVISION OF LABOR AND THE CONTROL OF PRODUCTION 75

Coöperation. The economic effects of division of labor. Division of labor and the extent of the market. Territorial division of labor. Division of labor a form of coöperation. The control of production. Regulation by legal authority. Protection for the individual. Promoting the national welfare. Mercantilism. Laissez faire. The progress of liberal ideas. The present system of economic freedom. Personal liberty. Private property. Individual initiative and control of enterprise. The motive to productive effort. What guides the entrepreneurs? The answer is price. What is back of price? Unconscious coöperation.

CHAPTER VI. THE FORM OF THE BUSINESS UNIT 94

The individual proprietorship. The partnership. The limited partnership. The corporation. The nature of the corporation. Capital stock. Bonds and notes. The property account of a corporation. Losses and insolvency. The value of stock. Interest payments and dividends. Stock dividends. Watered stock. The corporation as a factor in modern economic life. The marketing of corporate securities. Underwriting. Other marketing agencies and other functions. The stock exchanges. The services of investment bankers.

CHAPTER VII. LARGE SCALE PRODUCTION AND COMBINATION . . 120

The growth of large scale production. Conditions necessary for large scale production. Economics of large scale production. Power and machinery. By-products. Technical experiment and research. Economy in administration. The field of large scale production and its limits. Combination: the problem of large scale management. Horizontal combinations. Economy in buying. Economy in selling. Economies in operation. Saving in freight costs. Vertical combinations. Technical advantages. The United States Steel Corporation. Other examples of vertical combination. Conclusions. The weakness of combination.

CHAPTER VIII. THE ORGANIZATION OF MARKETING 139

Production to satisfy other people's wants. Bridging the gap between production and consumption. The function of trade. The evolution of exchange. Among primitive peoples. In civilized antiquity. In the Middle Ages. The traders. The town market. The medieval fair. Transition to the present time. Present-day importance of marketing. Marketing functions: assembling. Storing. The assumption of risk. Rearrangement. Selling. Transportation. Financing. Marketing agents. The retailer. The wholesale merchant. Other wholesale agencies. Markets. The produce exchange. Services of the middlemen. Evolution of marketing.

PAGE

Chapter IX. The Organization of Transportation . . . 164

Conditions before the nineteenth century. The era of road-building. The navigable rivers. The canal era. Development of the waterways and division of labor. The beginnings of the railroads, 1830–1850. The growth of the railways, 1850–1880. Government investment in railroads. Progress in organization, construction, and operation. Competition of rail and water routes. Railway abuses. The growth of the railways, 1880–1920. Economic services. Traffic on rivers and lakes. Highways. The merchant marine. Coastwise shipping.

Chapter X. The Economic Functions of Government . . 185

Government and the economic organization. Consumption of wealth. Production of wealth. Distribution of wealth. Varieties of government. Political organization in the United States. The economics of government. Defense. Justice and security. Regulation of industry. Protection against disease and accident. Care of dependents and defectives. Maintenance of moral standards. Protection against certain forces of nature. Protective and developmental functions. Education. Religion. Facilities for national industry. The primary functions of government. Government industries. Taking stock of our economic organization. Shall the system be scrapped? What the present system accomplishes. The present conclusion.

PART II. THE LAWS OF PRICE

Chapter XI. The Significance of Price. Demand . . . 217

The significance of price. Price controls the material enjoyments of life. Price controls trade. Wages and salaries are prices. Fluctuation in individual prices. Changes in the general price level. Importance of future prices. Complaints about prices. Should something be done about it? Is price a mystery? Prices fixed by statute law or custom. Legal price-fixing today. Customary price today. Neither sellers nor buyers are generally free to fix prices. Price not a matter of chance. A bargain sale. Price affects the amount of sales. The demand schedule. Market defined. The demand schedule is hypothetical. Time and place of the demand schedule. Purpose of demand schedule. Graphical representation. The demand curve. Smoothing the curve. Interpolation. Definition of demand. The relation between price and quantity chosen. The law of demand. Elasticity of purchase. Examples of inelastic and elastic purchase. "Necessities" and "luxuries."

Chapter XII. Supply 241

Response of sellers to price. Supply schedule and curve. Supply relates to a particular time and a particular place. Supply schedules and curves are hypothetical. The definition of supply. The relation

PAGE

between price and what sellers would offer. The law of supply. Elasticity of sale. Are the laws of demand and supply universal?

CHAPTER XIII. COMPETITIVE PRICES. THE DEFINITION OF DEMAND AND SUPPLY 248

Competition. A concrete case for illustration. The market price: quantities offered and taken are equal. The price at which exchanges take place. The price that clears the market. The case of the vertical demand curve. Generalizations. Price fixed by a sole buyer. Competitive price fixed by cost of production. Defining demand and supply. Two possible meanings of demand. Choice of a definition of demand. Two possible meanings of supply. Choice of a definition of supply. The importance of consistency. Demand and desire. Supply and stock.

CHAPTER XIV. THE CAUSES OF CHANGE IN COMPETITIVE PRICE . 263

What causes price changes? Effect of an increase in demand. Effect of a decrease in demand. Effect of an increase in supply. Effect of a decrease in supply. When the demand curve is vertical. When price is fixed by a sole buyer. When cost of production fixes price. Conclusion. When demand and supply both change. Increased demand; decreased supply: effect on price. Effect on quantity exchanged. Decreased demand: increased supply. When demand and supply both increase: effect on price. Effect on quantity exchanged. When demand and supply both decrease. Other cases of competitive price. Practical illustrations.

CHAPTER XV. THE LAWS OF COMPETITIVE PRICE: SUMMARY, QUALIFICATIONS, AND CONCLUSIONS 284

The laws of competitive price. Summary. "Theory and practice." "Market price" and the "perfect market." The "smooth" curve. Qualifications to the laws of price. Cause and effect in practice. The laws of price on the stock exchange. Shall we discard our theory? Purpose of this study. How to study problems of price.

CHAPTER XVI. WHAT IS BACK OF DEMAND? 296

Direct and indirect determinants of price. Total demand and individual demand. How utility varies. The subjective side of utility. Utility expressed quantitatively. Marginal utility. The curve of marginal utility. Meaning of marginal utility. The law of marginal utility. The psychological foundation. Regard for future wants. Marginal utility and total utility. Some apparent price anomalies resolved. Abundance and scarcity. Individual demand. Graphic illustration. Obtaining the maximum total utility. How both parties to a trade can gain. The decision to buy. Back to total demand. Real people. Is all buying irrational? Economic laws: the margin of

PAGE

error. The law of demand. Demand is responsive to countless influences.

Chapter XVII. Cost of Production, Supply, and Competitive Price 314

The different viewpoints of buyers and sellers. Business costs and the lower limit to price. The upper limit to price under competition. Price determined by cost. Prices not fixed by cost of existing goods. Future cost controls. Conclusion and qualifications. Analysis of cost. Relation between each producer's output and unit cost. The limit to reduction of unit cost. When the limit is reached. Relation between product and unit cost of a whole industry. Increasing costs. Constant costs. Decreasing costs. Marginal cost and average cost. Individual cost curve and supply curve. Supply and price: increasing costs. Constant costs. Decreasing costs. An unstable condition. Equilibrium. Qualifications. Cost of production and price changes. Competition. Combination and monopoly.

Chapter XVIII. Monopoly Prices 345

Classification of prices. Higgling prices. Monopoly. How the monopolist determines the price. Increasing costs. Decreasing costs. When the monopolist determines the quantity to be sold. Non-reproducible goods. High and low monopoly prices. Limitations upon monopoly price: substitutes. Potential competition. Risk of legal interference. Conclusion.

Chapter XIX. The Present Value of Future Income . . . 355

Income the goal of economic activity. Capital the means to the end. Property and capital. Property valuation. Present worth. Present goods more desirable than future goods. Interest. The discount process. Capital and income. Discount and the laws of price. Discounted income and cost of production.

PART III. MONEY, BANKING, AND EXCHANGE

Chapter XX. Money 371

Examples of primitive money. The common characteristic. Money is local. Definition of money. Functions of money. Gold and silver. Qualities of the ideal money commodity. Stability of value. Superiority of gold and silver. Government control of money. The monetary unit. Coinage. Free coinage. Making and regulation of paper money. Legal tender. Classification of money. The value of standard money. Essentials of standard money. The gold standard. Bimetallism. An example. Problems of bimetallism. Mint ratio and market ratio. When the two ratios are equal. When the market ratio is less than the mint ratio. When the market ratio is the greater. Gresham's law. The operation of Gresham's law takes

PAGE

time. Will a discrepancy between the ratios correct itself? The theoretical question. The practical answer.

Chapter XXI. Money (*Continued*) 391

Bimetallism in the United States. The original coinage ratio. The coinage ratio changed. The end of bimetallism. Change in the market ratio. European experience. Bimetallism in France. Conclusions regarding bimetallism. The "limping standard." The American silver dollar after 1873. Bland-Allison and Sherman acts. The result. A lost opportunity. Fiduciary money. Representative money. United States gold certificates. United States silver certificates. Token money. The value of token money. Quantity automatically regulated. Why token coins are light weight. Limited legal tender. Summary. Credit money. Irredeemable credit money. Fiat money. Experience with credit and fiat money. American money in the colonial and Revolutionary period. From 1789 to the Civil War. The greenbacks. An example of fiat money. Later history of the greenbacks. Credit money in the United States today.

Chapter XXII. The Principles of Banking 412

Financing production. Credit. Introducing the commercial bank. Example of the loan. Discount. Proceeds. Unsecured note. Secured note. Endorsed note. Draft or acceptance. A bank's balance sheet. A loan for cash. A loan for bank notes. A loan for deposits. A normal bank statement. Bank deposits. Discount and deposit. The check. Payment by check. Collection between banks. The clearing house. Clearing for outside banks. The volume of clearings. Out of town checks. Paying a loan. The normal business of banking. Function of the reserve. Profits *vs.* safety. How the reserve ratio is controlled. Buying and selling securities. Rediscounting. Control by the rate of discount. Note issue. Notes and deposits: fundamental similarity. Differences between notes and deposits. Legal restrictions. Bank credit as a medium of exchange. Currency. The need of elastic currency. Most money is relatively inelastic. Elasticity of bank credit. Influence of reserves. The note broker. The commercial paper houses. Trust companies. Savings banks. Private banks. The function of credit. The peculiar importance of bank credit.

Chapter XXIII. Banking Systems 446

Banking principles and banking systems. The French banking system. The Bank of France. British banking. The Bank of England. Concentration of reserves. Note issue. Elasticity is sacrificed. How Great Britain secures elasticity. Extreme procedure in time of crisis. The Banking Department. The reserve. The Canadian banking system. Notes. Few banks, but large ones. Colonial "banks." First Bank of the United States. Local banks, 1811–1816. The

PAGE

second Bank of the United States. Independent Treasury and local banks. The national banking system. The whole banking system. Security for national bank notes. Elasticity destroyed. Disturbing fluctuations in volume. Artificial value of government bonds. The illusion of bond security. A nation of small local banks. The need of elastic currency and centralized reserves. The federal reserve system. Organization. Functions. Rediscounting and discounting. Deposits. Federal reserve notes. Plan to retire the national bank notes. How the plan has worked. Centralized reserves. A national system of clearing. The government's financial business. Statement of the federal reserve banks. The United States monetary system.

CHAPTER XXIV. MONEY AND PRICES. THE VALUE OF MONEY . 476

Value of money not constant. What is the value of money? How the value of money is measured. Factors determining the value of money: volume of trade. Quantity of currency. Velocity of circulation. General conclusion. Standard or fiat money. Index numbers. Price relatives and the index number. The simple arithmetic mean. The weighted average. An index number of the aggregative type. Practical index numbers. Historical verification. An increase in the quantity of money: fiat money. A decrease in the quantity of money. An increase of trade. The end of fiat money. Increases in money and trade. Monetary inflation through gold. The world's production of gold. Effect on the price level. Monetary inflation of the World War. Soaring prices the result. Lower prices since 1920. Conclusion. Far-reaching effects of monetary fluctuations. Collapse of fiat money. The gold standard. Is a more stable standard possible?

CHAPTER XXV. THE BUSINESS CYCLE 502

The price level and the business cycle. Phases of the business cycle. The period of prosperity. Bank credit. Prices of securities. Checks to prosperity. The crisis or period of liquidation. Panics. The period of depression. The recovery. Motivating forces in the business cycle. Effects of fluctuation. Remedies for business fluctuations: control of credit. Forecasting. Stabilizing the monetary standard.

CHAPTER XXVI. INTERREGIONAL TRADE 520

Division of labor, trade, and transportation. The bases of interregional trade. Local advantages and their causes. The gain from specialization and trade. Conclusions. Diminishing returns and the law of comparative advantage. Transportation and the law of comparative advantage. The control of trade. Obstacles to trade and price equilibrium. The domestic trade of the United States. The

PAGE

foreign trade of the United States. The commodity character of our foreign trade.

CHAPTER XXVII. FOREIGN EXCHANGE 534

An export of cotton. The bill of exchange. Financing exports. Financing imports. The banks and the exchange market. The mint par of exchange. The rate of exchange. The gold points. Indirect exchange. Other types of bills of exchange. The exchange market. Fiat money.

CHAPTER XXVIII. THE INTERNATIONAL BALANCE OF PAYMENTS . 551

Mercantile fallacies. The balance of payments. Limits to the flow of gold. A gold producing country. Recent gold importations of the United States. Reaction of exchange rates upon imports and exports. The effects of borrowing and lending. The international balance sheet of the United States. The balance of trade once more. A favorable balance of trade does not mean that gold will flow in. Is gold a peculiarly desirable import? Is an adverse balance of trade a sign of national decline?

INDEX 569

PART I

THE ECONOMIC ORGANIZATION

ELEMENTARY ECONOMICS

CHAPTER I

FOUNDATIONS OF THE SCIENCE OF ECONOMICS

Among the various sciences which the pursuit of human knowledge has developed, there are those whose subject matter includes man himself, and among these are certain sciences which are concerned with the study of human beings, not as isolated specimens, as in physiology or human anatomy, but as members of social groups — families, tribes, villages, states, nations. These sciences having to do with man in his various relations with his fellow men are known as the social sciences, among which economics holds a prominent place.

The particular field of economics, which distinguishes it from the other social sciences, is derived from two of the most fundamental principles of nature, one having to do with the character of man himself, the other relating to the natural environment in which man lives. Let us seek to develop these two principles in order.

Human wants. Man is an animal, one of the millions of species of living things, plant and animal, that populate the earth. Nature has brought man and all these other forms of life to their present state as separate species and to their present equilibrium among themselves as sharers of the earth, through countless ages of evolution. The history of this process is a fascinating story, but one which enters only incidentally into the study of economics. Our task is to inquire into the nature of man as we find him in the world of today, without neglecting to give a nod of recognition to the fact that he is the product of a process of evolution, which is doubtless still going on, though so slowly that we are justified in considering man's present position in nature as fixed for all practical purposes.

The characteristic of man which most interests us at this point is the infinite number and variety of his wants, and the first stage in marking out the field of economics is to acquaint ourselves with the nature and peculiarities of human wants. For the first step or two we find nothing to distinguish man from the lower animals. In common with all species of animal life, man has certain organic needs which must be met or he will die. He must have food and water; he must have air to breathe. His bodily temperature must be always maintained somewhere between 98 and 99 degrees, and the remarkable thermostatic mechanism which nature has provided him for this purpose cannot cope with excessive exposure to heat or cold. Hence man requires certain forms of clothing and shelter, depending upon the climate. Man is also subject to attack by natural foes, and he must therefore have the means wherewith to defend himself or to escape. The simplest means appropriate to satisfy these fundamental organic needs are called "necessities," in the narrowest sense of that somewhat overworked term.

But now we come to a sharp distinction between man and the lower animals. The latter are satisfied when the necessities of life are met. With man the necessities are only the beginning. No sooner are they met than he reaches out for the satisfaction of other wants, for the " luxuries." Even the means whereby modern men satisfy the fundamental organic needs have passed far beyond the realm of bare necessities. This attribute of man is so fundamental to every phase of human life, and in particular to man's economic activities, that we must take pains to grasp its full import even at the risk of what may at first sight appear as an elaboration of the obvious.

Food. Food is necessary to prevent starvation, of course; but that is not what the average person is thinking about as he seats himself at the dinner table. He is thinking of the joy of eating. He studies the menu. He rejects oysters and decides on clams. He scans the soup list. He ponders on the choice of roasts:

chicken, turkey, beef, lamb, etc.; and so on down through the vegetables, salad, dessert, fruit, and coffee. This ceremony is something more than the warding off of starvation. And the same is true, only in lesser degree, of the simplest home table. To the original hunger instinct has been added a host of tastes and desires, so that we demand of food a good deal more than that it keep us alive. The infinite variety of desires that has grown up on the foundation of the drink-seeking instinct is too well known to need elaboration. Drinking is certainly not confined to the supply of moisture to keep the body alive.

Clothing. The necessity of protection from heat and cold gives rise to the desire for clothing and shelter. But to the modern civilized man or woman clothing means a good deal more than this. In buying a suit far more attention is given to the pattern, the cut, the style, the fit of the garments than to their cold-resisting qualities. Hats, coats, dresses, shoes, stockings, collars, ties, and so on down the list — how much of all this could be eliminated if all we wanted were protection against the elements! In fact the demand of civilized man for clothing is fully as much a response to vanity, and it is well known that in climates where clothing is not a necessity, such clothing as is worn by the savages is often the result of the desire for ornament to satisfy the wearer's vanity.

Shelter. In most climates some sort of shelter is a necessity. The Eskimo with his snow house and the tropical savage with his palm leaf thatch on a bamboo frame are catering to this elementary want and not much more. But the dwellings of civilized men, from ancient days down to the present, have been much more than this. From the humble cottage to the royal palace, men have sought to make of their dwellings, not mere shelters from the weather, but things of beauty, of comfort, of ostentatious display, of luxurious enjoyment. One will generally have to search long in the modern village or city to find any habitation that does not represent a craving for something beyond mere shelter. Our houses, like our clothes, are responsive to vanity. They also represent (often grotesquely enough, to be sure) our aesthetic desire for that which is beautiful.

Other wants. Thus we have commenced to list the wants of mankind. We might continue, mentioning for example some of the things which the need of defense requires: strongly built houses, locks on doors and windows, protective weapons, policemen, armies and navies and their equipment, drugs and medicines, hospitals, jails and insane asylums, insurance companies — this is simply the beginning of the list. Consideration of the gregarious tendencies of mankind reminds us of the concentration of people in cities, towns, and villages, thus giving rise to special wants in the field of housing, transportation, etc. Even when people seek relaxation from their ordinary activities and surroundings and take vacations "in the country," they do not generally seek solitude. They require seashore and mountain hotels, resorts, and camps, where they can take their vacations in company with others. Man has a thirst for knowledge, in its elements perhaps instinctive, but carrying him far beyond the promptings of mere instinctive curiosity. He wants books, telescopes, laboratories, schools, and colleges. He wants to travel and to visit new and strange scenes and so demands transportation facilities, hotels, and guides. We have a desire for the beautiful and so want music and works of art. We enjoy sport and recreation and therefore require golf clubs, baseball outfits, swimming suits, and all the paraphernalia of modern sport. We also like to take our exercise vicariously, and there must be baseball stands, race-track grandstands, ice rinks, football fields, etc.

Intermediate goods. We should also note here the fact that modern civilized man wants many things which are not capable of satisfying directly any desire. The housewife wants a sewing machine, though there is obviously no satisfaction directly obtainable from the possession or use of this article. But the sewing machine will serve as an instrument for making and repairing the family clothing, thus helping to satisfy a fundamental need. It is not wanted for itself, but as a means to the gratification of another need. This is a characteristic of the whole group of things known as instruments, tools, and machines, which make up a surprisingly large part of the commodities which men want and for which they strive. They are called "intermediate goods" and will come in for further study in the next chapter and elsewhere in our inquiry.

No limit to the wants of man. This sort of study might be carried to any length. Enough has been said however to bring out the important conclusion toward which we have been aiming; namely, the infinite number and the bewildering variety of the things that man wants. There is actually and literally no limit. Satisfying the instinctive cravings, securing the things necessary to sustain life and perpetuate the race — this is only the barest beginning. We soon forget that sustaining life is the basis of any of our wants. We are never content. The more our wants are satisfied, the wider is the horizon opening up new and tempting vistas of wants still to be catered to. We are quite justified in our conclusion that man's wants, as regards their number and variety, are without limit. This is the first of the two great principles which mark out the field of economics.

Nature, the provider. The second principle is that the wherewithal to satisfy man's unlimited wants does not come without effort on his part. Nature may be a bountiful provider; she is in fact the source of everything man has or can hope to have. Some of the things he wants she provides gratuitously. We have plenty of air to breathe; we have sunlight, though not always in the exact amount we might like; water is generally, though by no means always, furnished in abundance by nature. A certain amount of food, shelter, and clothing may sometimes be obtained from nature with little or no human effort, though this holds true only of the simplest wants of the most primitive men and even then only in exceptional cases. But after noting these few exceptions we must conclude that nature does not generally bestow her gifts with a free hand; practically none of the infinite number of things that man wants is furnished him freely by nature. Man must work for what he gets. Even the lower animals have to exert themselves to provide for their needs, and if man had been content with the sort of living which the lower animals thus obtain he would still be a lower animal himself. Man is always engaged in efforts to secure the things he needs for the satisfaction of his wants. His wants are unlimited; the things to satisfy them are strictly limited. There is never enough to go around. Here we have the explanation of the greater part of the activities of human beings through all history. Mankind has always been chiefly engaged in "working for his living." Here we have also the roots of the great human institutions, law, property, the family, etc., which have developed out of man's struggle for the things with which to satisfy his wants.

Economics. The insatiability of man and the niggardliness of nature are thus the foundation stones upon which rests the structure of economics. Starting with these two great premises, economics investigates the activities of man in his efforts to satisfy so far as he may his limitless wants out of the limited resources granted him by nature. We have then this preliminary and formal answer to the question: what is economics? *Economics is the science of man's activities devoted to obtaining the material means for the satisfaction of his wants.* Lest the reader miss the true inclusiveness of this definition, he is reminded of what will be made clearer a little later; *i.e.*, that the "material means" which serve to satisfy human wants include human beings as well as external objects.

When we stop to consider that an important part, not to say the greater part, of all the waking hours of most adult men and women is devoted in some form or other to the business of acquiring the means for satisfying wants, and that even the children are much engaged either in this same endeavor or in acquiring the education and training necessary to equip them to begin it a few years hence, we can realize that, although strictly limited by the definition, the field of economics is a broad one. We might proceed to catalogue the many topics included and the various problems to be investigated. But a mere glance at the table of contents of this book will perhaps suffice to give the reader a general notion of the subject matter of economics sufficient for the present to indicate the importance of the subject and the advantage to be gained from its study. We may let this stand as the preliminary answer to the demand for a notion of the content of economics and the reason for its study. The complete answer may be expected to unfold itself as our inquiry proceeds.

Economics as a science. We have been speaking of the science of economics, and it may help the reader to a still clearer notion of what economics is and also enable him to avoid some seductive pitfalls as he pursues his further path of investigation if we pause here to inquire briefly into the nature of science and the distinction between science and art. We cannot do better than quote the following words of a distinguished American economist of the older school.[1]

"A science, whether the science of mathematics, or physics, or mechanics, or chemistry, or geology, or physiology, or economics, deals only with the relations of cause and effect within its own field. It does not start with the notion that something is desirable or undesirable; nor does it arrive at any such conclusion as its result. It has no business to offer precepts or prescriptions. Its sole single concern is to trace effects back to their causes; to project causes forward to their effects.

"An art, on the other hand, starts with the assumption that a certain thing is desirable or that a certain other thing is undesirable; that something is a good or that something is an evil. The object it seeks is to ascertain how the good may be attained, or the evil avoided. In pursuing this inquiry, it makes use of the principles, or laws, governing the relations of cause and effect, which have been ascertained in the cultivation of any and all sciences that have in any way to do with its own subject matter. As a result, it issues with certain precepts and prescriptions for the guidance and assistance of those who would gain the good or avoid the evil which that particular art has in contemplation, whether it be the art of navigation, or of cookery, of painting, of gunnery, of architecture, of mining, or of weaving.

"This distinction between a science and an art ought to be sufficiently clear; but the inveterate disposition of economic writers [to confuse the two] which has been referred to, will perhaps justify an illustration which I shall make familiar, even at the risk of appearing coarse.

---

[1] Francis A. Walker, *Political Economy*, Third Edition, pp. 19–21.

"Suppose I am in my laboratory and a man enters who says that he desires to consult me, as a professor of chemistry, as to whether he had better swallow the contents of a vial which he holds in his hand. I reply to him: 'Sir, I have no advice, as a professor of chemistry, to offer you as to what you shall swallow or refrain from swallowing. I perceive that the liquid contained in your vial is prussic acid. I will cheerfully state to you the action of prussic acid on any substance about which you may choose to inquire; but probably you had better, for your apparent purpose, go to Prof. S., the physiologist, who can more fully and readily than myself explain the precise action of prussic acid when taken into the stomach of a living being.'

"The inquirer now goes to Prof. S., and says that he desires to consult him, as a professor of physiology, as to whether he had better swallow the liquid which the chemist has told him is undilute prussic acid. Prof. S. replies: 'Sir, should you consult me as a fellow being, I would not stand on ceremony, but frankly advise you to empty the contents of your vial into the sink. But if you insist on consulting me as a professor of physiology, I must reply that I have no advice to give. Physiology, sir, is a science; as such, it has nothing to do with precepts or prescriptions, but only with the relations of cause and effect within the field of animal life. As a student of that science, I inform you that, if you swallow the liquid, you will experience such and such sensations, and, at about such a time, you will be dead. Since you still insist upon having advice as to whether you had better do this or not, I refer you to my neighbor, Dr. G., who is the professor, not of a science, but of an art. As such, it is his business to give advice regarding conduct. As such, he has a right to entertain the notion that certain things are good, and certain things evil; that the means calculated (as shown by the appropriate science or sciences) to bring about the good, are desirable; that the courses which (as shown by the appropriate science or sciences) lead to evil, are undesirable. He would not be a physician unless he held that pain and death were evil; life and the absence of pain, good. What he is a physician for is to help his patients to avoid the evil and obtain the good. In doing this he will naturally seek to apply the largest and latest results of the *science* of physiology to the *art* of healing.' "

In thus sharply marking out the field of science we do not imply that moral or ethical considerations are to be ignored by the student of economics. Economics is a social science, and as such it grasps as its subject matter all attributes of man which affect his pursuit of the material means for the satisfaction of his wants. Among these attributes the moral and ethical notions of mankind are of the utmost importance and require full consideration. For example, the moral code of one community may prescribe polygamy, while that of another insists upon monogamy, and these moral edicts are of great economic significance. Only we must not forget that the science of economics can not and does not attempt to tell us whether monogamy or polygamy is right or wrong.

Utility. We need now to acquire clear notions of certain fundamental concepts which are inherent in the subject matter of economics and are an essential part of the equipment of the student. Man desires those things that are able to satisfy his wants. Anything that satisfies a human want is useful; it has utility. *Utility is that quality of a thing which enables it to satisfy a human want.* This definition scarcely needs elaboration or explanation. We determine readily enough whether we want any particular thing; *i.e.,* whether it is " useful " or " useless."

The economist gives somewhat broader meanings to " useful " and " utility " than are employed in common speech. When we call anything useful we imply nothing as to its merits in an ethical sense. A thing is useful if it can satisfy a human want; that is all we ask. We do not inquire whether the satisfaction of that particular want may be good or bad, wise or foolish. Bootleg whisky is in this sense of the term useful. It may be morally bad, as an intoxicant, it may even be rank poison, but such facts do not alter its classification as an object of utility so long as it does actually enable a human want to be satisfied. The burglar's jimmy satisfies a want of his; it has utility in the economic sense.

The use of chewing gum is probably foolish, but we cannot thereby deny its utility. The pictures which adorn the walls of many homes are probably wholly bad from an artistic sense, but they are nevertheless economically useful. Of course this does not imply any lack of appreciation of the importance of aesthetic and ethical judgments. Such questions are simply outside the field of economics, as of any other science.

Definition of wealth: Utility. If from the whole category of useful material things we eliminate those few which nature furnishes man in abundance sufficient to satisfy his wants, we shall approach what the economist calls "wealth." The science of economics is largely built up about wealth, so much so that economics has often been defined as "the science of wealth," a definition which, properly understood, is almost synonymous with that which we have accepted. It is indispensable therefore to have a clear idea and a precise definition of wealth.

Wealth consists of all useful material things owned by human beings. This definition needs some further examination. We have already settled the meaning of "useful." Useless things are not included in wealth, which thus embraces only those things which are capable of satisfying human wants. The reader should perhaps again be warned that the ethical character of human wants does not enter into the classification.

Materiality. We limit wealth secondly to material things. This excludes such desirable things as honesty, good health, the speed of a race horse, the coldness of ice, the skill of an artist or a musician. This may possibly seem arbitrary and unfortunate at first. But we shall not have to look far to see the great advantage of this limitation and particularly the confusion which would follow if the concept of wealth were not thus limited. If the speed of the great horse Zev is wealth, what of the horse himself? Certainly he is wealth. But have we then two articles of wealth, the horse and his speed? And if so, should we not add a third, his strength; and a fourth, his beauty? Most of the forms of immaterial wealth with which we should have to deal if our definition admitted these are simply qualities of material things which are themselves wealth. To count both the material thing and its qualities as wealth would be the crudest sort of double counting and productive of nothing but confusion. This is really to confuse wealth and utility, since most of the examples of immaterial wealth which would be urged are simply those qualities of some material object which enable it to satisfy a human want. Similar confusion would arise from including human qualities, such as health, strength, skill, honesty, etc., in the category of wealth. We shall consider a little later the question whether human beings themselves are wealth.

Wealth must be owned. Having excluded from our concept of wealth all useless and all immaterial things, we next limit it to those useful, material things which are owned by man. The first effect of this limitation is to exclude those useful things which are furnished freely and abundantly by nature. Thus air is material, and it is about the most useful thing in the world. But on account of its abundance it is not owned. Why should one take the trouble to own it? We also exclude certain things which, from their nature or from the nature of the human institution of ownership, cannot be owned. We cannot own the sun or the moon; no person or group of persons can own the Atlantic Ocean or the Gulf Stream. Most material things that satisfy human wants however are owned. Ownership may be vested in the individual person, as in the case of one's suit of clothes, the farmer's land, the merchant's store, etc. Or various groups of persons, such as partnerships, corporations, etc., may own land and buildings and virtually any kind of wealth which may be owned by individuals. Governments, national, state, and local, are large owners of wealth; they are simply another form of association of human beings.

Are human beings wealth? The requirement of ownership by human beings restricts the scientific concept of wealth in another important respect. Are human beings themselves wealth? As regards slaves, the question is clearly answered in the affirmative by reference to our definition. The slave is useful; he is a material thing; he is owned by his master. But how about free men or women or children? They are not owned under the legal systems of most modern nations. We might set up the fiction that free persons are owned by themselves or (in the case of minor children) by their parents. Free persons would then fall into our category of wealth, and there are some decided advantages to this classification. Free persons are material things, of course, and they have qualities which enable them to satisfy wants of other persons. Moreover it is not always easy to know just where to draw the line between free men and those who are owned. Logic and simplicity in the analysis of wealth would perhaps be promoted by including all human beings. On the other hand it is not customary to think of free persons as owned, even by themselves, nor do we naturally think of free persons as wealth. The weight of advantage is on the side of our definition as stated, according to which, while slaves are undoubtedly wealth, free persons are not. Throughout the whole study of economics however it will be found necessary to take account of the utility of free persons and their services in satisfying human wants on the same terms as the utility and services of wealth.

How to identify wealth. We thus arrive at our definition of wealth. To determine whether any given thing is wealth or not, we have only to ask three questions: is it useful? is it material? is it owned? If the answer is yes to all these questions, the thing is wealth. A negative answer to any one of the questions excludes it.

The problem of definition. This method of setting up the definition of wealth may possibly appear unnecessary and arbitrary. Do we need a precise definition of such a common word as wealth? And if so, have we any right thus to rig up a definition to suit ourselves? In answer to the first question, we need to remind ourselves that agreement upon precise definitions of all technical terms is one of the first requirements of science. Economists have not always realized the importance of this requirement or that it applies to their own science no less than to others. It is sometimes assumed that, since the technical terms of economics are largely those of everyday speech and writing, everyone must have a clear enough knowledge of their meanings. It is agreed that the physicist must define "ampere" and "volt," because no one would otherwise know what he was talking about. But why should the economist have to define "wealth," "income," "money," and other common terms? Doesn't everyone understand well enough what these words mean?

Now the fact is that everyone does not understand well enough what these terms mean. Ask a dozen intelligent persons to define wealth; you will probably get a dozen different definitions. Is sunshine wealth? Is a burglar's jimmy wealth? How about the moon, the skill of a great surgeon, the town hall, the month of January, the Gulf Stream, the honesty of a bank cashier, etc.? People may think they have clear ideas of the meanings of common words until they are asked to formulate and defend a definition; then the vagueness of popular usage becomes evident. Such vague ideas may do well enough for ordinary use, but science requires precision. Economics therefore cannot evade the obligation of defining its technical terms, even though they be words of everyday use.

Indeed the adoption of such common terms makes the problem of definition harder rather than easier. Everybody will admit that there must be a definition of "ampere" and will see that there is nothing to do but learn and use the definition exactly as given. But economics must always contend against a certain feeling that definitions of common terms are of no great importance and against the vague and contradictory preconceptions which have become attached to the words of everyday use. This difficulty must be met squarely; it cannot be evaded.

Having accepted the obligation to define our technical terms, we revert to the question: are we free to set up our definitions as we see fit? The answer is yes. Names and definitions are to a certain extent arbitrary. There is no absolute test of the correctness of a definition; it is idle to argue about it. The reader who is familiar with Mark Twain's compilation of *Extracts from Adam's Diary* will recall the frequent disagreements that arose between Adam and Eve over the naming of the animals in the Garden of Eden, a task which Adam found sorely taxing his powers but which was simplicity itself to Eve. When for the first time there appeared the dodo, Adam was at a loss for its name. Eve announced the correct name at once and, to all demands for proof or explanation, replied scornfully " the moment one looks at it one sees at a glance that it looks like a dodo." So it became the dodo. If in the development of the English language the term " dog " had become attached to the biped now known as a duck and the term " duck " had similarly become associated with our favorite quadruped friend, we should get along just as well as we do now. Words have meanings which have become attached to them through the long process of the evolution of language. To argue whether these meanings are correct or not would be a foolish waste of breath. And nothing is to be gained by an argument over the " correct " definition of wealth.

The requirements of a scientific definition. On the other hand, definitions are not a matter of indifference. In science much depends upon the choice of terms and their careful definition. While the test of absolute correctness is wanting, there are two simple pragmatic requirements to which every scientific definition should conform. First the definition must be useful for the purposes of scientific study, and second it should agree as closely as possible with popular usage. These two requirements may come into conflict with each other, and in that case it is generally the second which must give way. This sacrifice is not so serious as might at first appear. Since popular usage is always more or less vague, a scientific definition must in any event differ somewhat from the popular definition or definitions. Exact agreement with popular use being out of the question anyway, the departure required to make the definition useful is generally not serious.

We justify our definition of wealth, therefore, on the ground that it will be useful in our further investigation without departing too far from common usage. That we have fairly well met the second of these tests will probably be apparent at once. The usefulness of our definition will not be fully demonstrated of course until the student has finished this book, or better still after he has continued his study of economics with the aid of other authors or through his own research.

Income. Wealth must be useful. When an article of wealth satisfies a want of man we say that it furnishes *benefits* or renders *services.* By the benefits or services of wealth we mean *desirable events which it causes for human beings.* Keeping out the cold is a desirable event caused by a dwelling house; it is one element of the service of shelter or benefit yielded by the house. The service of a piece of ice in the refrigerator is keeping things cold. A diamond ring renders service by satisfying the love of beauty or perhaps only the vanity of its wearer. The service of a slave is working in the field. Thus every article of wealth renders its appropriate service to man, else it would not be useful and so would not be wealth. In exactly the same way services are rendered by free persons, as when an actor performs before an audience or a lawyer tries a case for his client. The word *income* is used to include such services and is defined as follows: *Income consists of the benefits or services rendered by wealth or by free persons.*

Costs. All the events caused by wealth are not desirable. Besides furnishing shelter, a dwelling house requires painting and repairs, it compels its owner to pay taxes and insurance premiums, and so on; finally it wears out or becomes obsolete and if not discarded must be torn down and rebuilt. The slave works in the field, but he requires food, clothing, and shelter, and the free farm laborer requires wages. *The undesirable events caused by wealth are called the disservices or costs of wealth.*

Net income. All wealth yields both income and costs. We cannot escape the undesirable events if we would have the desirable events; we put up with the costs in order to get the income of wealth. It follows that, in the long run, the income from any article of wealth must exceed its costs (how these are measured will appear later), else no one would care to own that article and it would cease to be wealth. For example a storekeeper has a motor truck to deliver goods to his customers. The truck is an article of wealth; the income from it is the service of delivering goods; its costs are the expenditures of money for the gasoline and oil it consumes, garage room, occasional washing, and repairs. These costs would not be borne if the service of delivering goods were not worth more than the costs. Eventually the truck begins to wear out; its speed diminishes and it cannot make some of the grades. It consumes more gasoline, and its loose piston rings and bearings are wasteful of oil. Repair bills become heavier and more frequent. Thus its income declines while its costs rise, and sooner or later the margin between them will disappear. The machine is no longer useful and it finds a resting place on the junk heap; it is no longer wealth. *The difference between the income and the costs of any article of wealth is its net income.* Some net income must be present or the article will not be wanted. In general then utility means the power to yield a net income or an income in excess of the costs, and this idea of a net income is inherent in the term "useful" as employed in the definition of wealth.[1]

Property. Every article of wealth is owned by somebody. This relation between wealth and its owner is one of the most fundamental concepts of economics. It may be called ownership, property right, or property. *Property* is the technical term generally used in economics. It is defined as *the right to income; that is, the right to the benefits or services of wealth or free persons.* This evidently is what ownership means. When a person owns an article, he has a right in it; namely, the right to have it and use it, all of which is included in the right to the services or income from that article. The owner is protected in his property right by the laws and customs of the community in which he lives; other persons are prevented from interfering with his enjoyment of the services of his wealth.

Division of property. A sole owner has the exclusive right to all the income of the article of wealth in question, as in the case of the owner of a house and lot, an automobile, a suit of clothes, etc. Ownership may however be divided, in which case each of the two or more owners has certain rights in the wealth;

---

[1] The reader who has any familiarity with accounting will note that these definitions of income, costs, and net income differ somewhat from those used by the accountants. They are not really in conflict. The accountant's terminology represents a special technical application of the general concepts with which we are here dealing.

that is, the right to certain of its benefits or income. The tenant of a rented house has the right to its use only during a certain time, as limited by his lease contract with the landlord, who has retained the right to the services of the house at all other times; that is, before the beginning and after the termination of the lease. In a partnership the wealth held belongs jointly to the two or more partners. Each has the right to a certain part of the services or income of this wealth, the exact division being specified in the partnership agreement. So in the corporate form of business organization, the wealth of the corporation belongs ultimately to the stockholders, of whom there may be hundreds or thousands, each with certain rights in the income from the corporation's property. The bondholders of the corporation represent another group of part owners, with certain other rights in the corporation's wealth, all specified in the bonds. Thus each bondholder has the right to receive certain interest payments at certain specified dates and finally at a specified date to receive a certain larger sum equal usually to the amount which was loaned to the corporation. Each stockholder has the right to a certain fractional part of the income of the corporation's wealth which remains after the rights of bondholders and possibly other classes of part owners, such as note holders, etc., have been satisfied. Property may thus be divided in a great variety of ways. Every person who has a property right in an article of wealth is an owner (sole or part) of that wealth, though in common speech the term "owner" is sometimes applied to only one of the parties. Thus the landlord and the tenant are really both owners of the rented house. Neither is the exclusive owner, for neither has the right to all the services of the house. The property is divided between them, though we are accustomed to apply the term "owner" to the landlord alone. In like manner the bondholders are not usually spoken of as owners of the corporation, though fundamentally they are part owners as well as the stockholders.

All the services of wealth must belong to somebody. When property is divided the sum of all the property rights of the several part owners must equal the total property right to all the services of that piece of wealth. For every article of wealth there must be a property right or property rights, and back of every property right there must always be an article of wealth or a free person or free persons.

Property in free persons. The last words of the previous paragraph will recall that property may be a right to the services of free persons as well as of wealth. If we had chosen to include free persons in the definition of wealth, all that has been said of property in wealth would have applied equally to them. As it is we must recognize that free persons render services to others and that these others have rights to such services. A baseball player signs a contract with a manager agreeing to play exclusively for him during the coming season. The manager has thereby a property right, the right to receive certain services from this player. To that extent the player's right to his own time and services — actually his right to himself — is diminished. It is a case of divided property again. Theatrical managers, " movie " producers, musical promoters, and others have property rights in actors and actresses, musicians, and other artists. Every employer has a property right to certain services of his employees. A promissory note gives the creditor a property right in the debtor; that is, the right to receive a certain sum of money. Lawyers have rights against their clients, and clients against their lawyers; parents have rights in their children, and children in their parents. Every business contract gives each party a property right against the other. In fact the property in persons is almost always divided. Practically no one is completely his own master, and there are all degrees of property rights of others, from the slave, a true article of wealth because completely owned by his master, through serfs, peons, indentured servants, to the freest so-called " free man."

Wealth and property. In popular speech, the terms property and wealth are not carefully distinguished. It is quite usual for example to speak of a house and lot as " property," to call household furniture " personal property," and so on. For the purposes of scientific analysis, economics draws a sharp distinction between property and wealth. The latter is a physical, tangible thing; it is material. Property, on the other hand, is a relation between wealth (or free persons) and persons; it is immaterial. This distinction is fundamental and has many important bearings. Thus wealth is the source of income; whereas property is the distributor of income. The wealth of a community, together with its productive free persons, determines how much and what kind of income the people may enjoy. The property rights of the several members of the community determine how these incomes shall be divided among them, how much of the total community income each one shall enjoy. Property rights may be changed without changing the total amount of wealth. The material welfare of the human race is a matter of the satisfaction of wants and is affected both by the total quantity of wealth and by the way the property rights to wealth are distributed among individuals, families, and nations.

Documents in evidence of property. Property rights are sometimes, though by no means always, expressed in legal documents. The owner of a farm or a building lot has a *deed,* which is a piece of paper containing a legal statement of his property right to all the benefits of a specified parcel of land. When one purchases the right to sit in a certain seat of a theatre on a certain date his property is evidenced by a ticket. On the other hand, no legal document testifies to one's property right in the clothes he wears or the food on his table. Whether there shall be a document or not is a matter of custom or law. The property is the original and fundamental thing; the document, when there is one, is secondary and incidental, and in many cases no document is considered essential. The reader must be on his guard against confusion of property rights and the documents in evidence thereof, particularly in those cases where common, and even scientific, usage has not found it necessary to employ different terms for these two things. For example a bank note is, strictly speaking, a piece of paper testifying to a property right of the bearer against a bank. Yet it is common to speak of the bank note as though it were the right itself. When once he has been warned, the reader is not likely to have any difficulty over such usage, loose though it confessedly is.

Wealth and income: fund and flow. There is a fundamental difference between wealth and income, which manifests itself in the matter of measurement. When wealth is measured it is the quantity in existence at a certain time that is determined. In measuring income we ascertain the amount of service that is rendered during a certain period of time. We do not speak of the yield of an orchard or the service of a carpenter at 1.25 P.M. of a certain date. It is ordinarily impossible to think of income except in reference to an appreciable length of time. An orchard yields so many bushels of apples per year, a mason lays so many bricks per day or per hour, etc. Wealth is a *fund;* it is measured as of a certain *instant* in time. Income is a *flow;* it is measured as of a certain *period* of time.[1]

Value as a common denominator. Each kind of wealth, income, or property has its appropriate unit of measure. But there is one unit in which all kinds of wealth, income, and property may be measured and which is by far the most important of all units. This is the unit of value or price. The universality of exchange is one of the characteristics of the modern economic organization, to which much attention will be given in the succeeding pages. Value and price are, next to wealth itself, perhaps the most important of the fundamental concepts with which the student of economics must equip himself at the outset of his study. We must examine this subject closely and formulate our notions with clearness and precision, perhaps at the risk of introducing once more some fairly obvious concepts.

Transfer and exchange. When Mr. Jones gives his dwelling house to his wife, there is a change of ownership. The house, which formerly belonged to Mr. Jones, now belongs to Mrs. Jones. A *change of ownership of wealth* is called in economic terminology a *transfer*. Transfer takes place whenever wealth is bought or sold, bequeathed, or given away. Now suppose Mrs. Jones sells the house to the Brown Realty Co. for $20,000. There has been another transfer of the house. But this time there has been also a transfer of money, from the Brown Realty Co. to Mrs. Jones. These two transfers together make an *exchange*, which is defined in technical language as *a pair of voluntary transfers between two owners, when each transfer is made in consideration of the other.* A gift is not an exchange, since there is only one transfer. As is said, there is " no consideration."

---

[1] By the methods of the infinitesimal calculus, the period taken for the measurement of income may be reduced to zero, thus making possible the measurement of income as a *rate* of flow at an instant of time. This concept is of great importance in certain advanced problems of economics, as it is in some of the natural sciences. It need not further concern us in this elementary book.

Barter and money exchange. *Exchanges made without the use of money are called barter.* For example the farmer brings ten dozen eggs to the village storekeeper and takes in return ten gallons of gasoline and ten quarts of motor oil. This transaction fits the definition of an exchange. " Swapping " horses or anything else is barter. At some time in the history of any people, barter was probably the usual form of exchange, such as there was, and a certain amount of barter always goes on among all peoples today. But the civilized world long ago abandoned barter as the normal form of exchange. In practically all exchanges today one of the transfers is a transfer of money, and money exchange has thus taken the place of barter. The subject of money will come up for special study in later chapters. For the present it is sufficient to know that *money consists of articles of wealth and property rights which are generally accepted in exchange for other wealth and services.* The money of the United States, as everyone knows, consists of gold and silver coins, various forms of " paper money," and small-change coins of nickel and copper. The unit is the dollar.

Value. From exchange come two of the most important concepts of economics and business, not to say of modern human life in general; namely, value and price. As the term is employed in economics, *the value of anything is the quantity of any other thing that would be given in exchange for the first thing.* If ten dozen eggs are exchanged for two bushels of potatoes, the value of the eggs is two bushels of potatoes, and the value of the potatoes is ten dozen eggs. If a ton of stove coal is sold for twelve dollars, the value of the coal is twelve dollars, and the value of twelve dollars is one ton of stove coal. It is not customary however to speak of the value of money in this way. On the other hand values of practically all other things are regularly expressed in terms of money.

Price. *The price of anything is the amount of money that will be given in exchange for one unit of it.* If a pound of cotton will sell for twenty cents, the price of cotton is twenty cents per pound. In stating the price of anything, the unit of measure used must always be stated or understood.

Value and price. The terms value and price are of course used more or less vaguely and inconsistently in popular speech, and there is some difference in usage among economists. The definitions laid down here agree as closely as is possible with popular usage, and they will prove themselves the most useful ones for economic analysis. As here defined, value and price are closely related to each other. They are distinguished from each other in that (1) value may be expressed in terms of any kind of wealth, property, or service, while the price of anything is always expressed in money, and (2) the term value is used for any quantity of a good, whereas price relates only to one unit. In practice value as well as price is usually expressed in terms of money; in which case value is price multiplied by quantity, and the value of one unit is identical with the price. Thus if a farm sells for $25,000, we may say that the value of the farm is $25,000, or that its price is $25,000. We should not say that the price of fifteen tons of coal is $150, but should say that the value of fifteen tons of coal is $150, because the price is ten dollars a ton.

The unit of value. Value thus gives us a common unit in which to measure all kinds of wealth, income, and property. This is the unit of money value; *i.e.*, the value of the monetary unit, in America the dollar. Since value is almost always expressed in terms of money, we may ordinarily call the dollar the unit of value, not bothering always to say the unit of "money value." Thanks to this common unit, it is possible to total a list of various articles of wealth and property rights. Bushels of potatoes, yards of cloth, pounds of sugar, acres of land cannot be added together. But the values of all these things can be. It is possible therefore for the business man to make an "inventory," for the citizen to make up his "tax list," for anybody to make a list of his possessions and get the total. Above all, it is possible to keep accounts of wealth and property and income.

EXERCISES

1. Which of the following may properly be called wealth? Give a reason in each case.

   (*a*) A gold coin. (*b*) Gold ore in a mine. (*c*) Gold in the sun. (*d*) An autograph of Benjamin Franklin. (*e*) A healthful climate. (*f*) Executive ability. (*g*) A farm whose ownership is under dispute.

2. According to definition did land on Robinson Crusoe's island have a value?

3. What benefit is rendered by:

   (*a*) A pound of sugar. (*b*) A pleasure automobile. (*c*) An acre of wheat land. (*d*) A customs house inspector. (*e*) Whisky containing wood alcohol. (*f*) A war medal. (*g*) A roulette wheel. (*h*) A moving picture actor.

4. Are the following property? If so, on what is the property right based?

   (*a*) Life insurance policy. (*b*) Bank note. (*c*) A signed promise to subscribe for the *News*. (*d*) Ticket to a football game. (*e*) Building site. (*f*) A gambling obligation.

5. During the war the United States borrowed billions from the people through the sale of Liberty Bonds.

   (*a*) Does this debt affect the wealth of the United States?

   (*b*) During the years these bonds are outstanding, interest will be paid the bondholders out of funds raised by taxation. Does the annual payment of interest on a national debt of this kind affect the wealth of the country?

   (*c*) Will the wealth of the country be affected if the bonds are paid off?

NOTE. This exercise refers to immediate effects, not to remote or indirect consequences.

utilities
form
time
place

CHAPTER II

THE FACTORS OF PRODUCTION

Production. Our first step in the study of production must be to settle exactly what we mean by the term. Who are the producers of the world? Shall all who engage in any form of mental or physical labor be considered producers, or is there a dividing line; are some engaged in productive and others in unproductive labor?

Common usage would lead us to say at once that the farmer, who raises wheat, is a producer; he obviously produces wealth, in that he creates something which did not previously exist. But a little reflection will show that he is a creator or producer only in the sense that he has been instrumental in effecting a change in the form or relationship of various physical elements already in existence so as to make them suitable for human consumption. The cotton manufacturer likewise adds nothing to the world's store of materials, but when he twists cotton fibre into thread and weaves thread into cloth he has added to the utility of the cotton in the eyes of the consumer. If we examine any other form of agriculture or manufacturing, we shall find that the essence of the particular operation is to bring about a change in the form of the good which will bring it one stage nearer the condition where it may satisfy a human want. The production or creation of wealth means then a change which renders the article more useful; in other words, it is the *creation of utility.*

Thus far we have considered only one of the many phases of business activity. There seems to be little similarity between farming and the operation of a railroad if one views only the technique of each business, but the purpose is essentially the same. Wheat on the farms of the West is quite as useless to the consumer of the East as if it had never been grown, and the service of the railroad in bringing it to the consumer is just as essential as that of the farmer in growing it. The grocer who keeps flour in stock to meet his customers' demands is performing a similar service, for the consumer not only wants a particular good at a particular place, but he wants it at a definite time. Today's flour is patently more useful in satisfying today's hunger than flour which will not be available until next week.

Thus are illustrated the three forms which the production of wealth may take — the creation of *form utility,* of *place utility,* and of *time utility.* All effort directed toward the production of wealth may be classified under one or another of these three headings; agriculture and manufacturing result in the creation of form utility, all the varied transportation services are engaged in creating place utility, and merchants of all classes are creating time utility. To the first one alone the name production is sometimes given, doubtless because of the obvious character of the change, but in view of the essential nature of the other two forms it is difficult to justify such usage, and we shall use the term production of wealth to include all three.

Although our primary interest in the study of production centres around the production of wealth, it should not be forgotten that man has wants which cannot be satisfied by wealth alone. A mechanical piano may satisfy the longing of one for music, but to another it is a sorry substitute for the performance of the pianist. If the labor of the pianist in making the roll for the mechanical piano is productive, certainly his performance on the concert platform must also be productive. We must therefore include in our definition of production the rendering of direct personal services by free persons. Completely stated, *production consists of* (*a*) *the production of wealth, which is defined as the creation of utility in wealth, and* (*b*) *the rendering by free persons of direct personal services.*

This definition is intentionally broad, including any effort which results in the satisfaction of a human want. In defining utility, in Chapter I, it was pointed out that scientific definitions have nothing to do with questions of ethics or morals. Like utility, the word production carries with it no moral significance, although economists are not therefore unaware that effort spent in some directions may be less commendable than effort spent in others, and that the results of human exertion may be fatal to individuals or disastrous to society. The study of production is the study of human activity in satisfying wants, whether or not that activity can be approved on moral or ethical grounds.

The factors of production. In spite of the fact that man, physically and intellectually, is today probably little different from what he was at the time when historical records begin, and in spite of the fact that the natural forces and laws have remained the same, man's efforts devoted to the satisfaction of his wants have during the period covered by history immeasurably increased their effectiveness. The explanation lies in the fact that man has gradually accumulated knowledge of the natural environment and learned to make it serve him in production.

This power to guide and control natural forces to his own ends marks one great difference between man and the beast. Animals take what is offered them, and if their necessities are not supplied, some or all of them must die. Man struggles so to change the environment that there shall be enough of the necessities to keep him alive; he forces nature to produce the plants and animals he wants. He has also learned how to make tools and machines which make his own labor more effective. Centuries before the beginning of historical times man must have roamed the fields seeking wild berries and edible roots for his food, and his early attempts to produce food were crude in the extreme. But with centuries of striving improvements have come, so that at present the same effort results in many fold the product of former times. The economic history of mankind is, from one point of view, the story of how men have secured food, shelter, clothing, and the comforts and refinements of life in increasing abundance and with less and less effort.

Man thus enlists in his service two powerful partners, and these three — man, nature, and capital — work together in production, with results which, when compared with man's first efforts, are almost unbelievable. These three partners are commonly called the *factors of production* under the titles, *land, labor,* and *capital,* the first two being figuratively used to stand for nature (not merely land in the narrow sense) and man. The rest of this chapter will be devoted to the study of their place in the economic organization.

Land. Land is essential in providing " standing room " for the population of the earth; it provides the food supply for the people and the raw materials for the factories and is the original source of all that man has. It is more than mere land, for it embraces the whole natural environment, including the oceans, lakes, and rivers, the topography of the land, whether a level or a mountainous country, the character of the soil, the amount and seasons of the rainfall, the range of temperature at different seasons, the mineral deposits, the water power, and a host of other features, all of which are important and may have a determining effect on the character of occupation and the type of industry in a particular region.

There are many examples illustrating the influence of the character of land on the industrial life of a locality. The conjunction of coal and iron in close proximity to each other almost inevitably stimulates the building of metallurgic works and leads to the growth of manufacturing; the point where one large river empties into another invites the building of a trading centre; the even humidity of the atmosphere may give a region an advantage in the making of fine cotton cloth; the presence of waterfalls in the past invited the founding of gristmills or cotton mills and at present determines the location of hydro-electric plants. While many industries are located in places which seem at the present time to offer few natural advantages, investigation of the cause of the original establishment will usually disclose that there was at one time some real advantage. At the time when cotton mills were first started in the United States the only available power was furnished by running streams. Hence the mills were located in places where water power was available, such as Lowell or Lawrence. Today water power is not of such importance as steam power, yet we find the textile industry of Massachusetts still largely concentrated in the region of its birth. The reason is that when an industry is once located in a place it attracts other manufacturers in the same line of industry. There may no longer be natural advantage in the locality, but there are artificial advantages, such as a labor force skilled in the particular trade, the presence of subsidiary industries which can take care of minor operations, transportation, banking, and other facilities devoted to that particular trade.

Labor. Land without labor is useless for the purposes of man. Even in the spots most favored by nature some effort is required to secure the goods desired, and on most of the earth's surface much exertion is necessary to make a living. For the majority of men labor is painful and is avoided if possible. Only when the desirability of things which can be secured by labor is considered to outweigh the unpleasantness of mental or physical exertion will the effort be undertaken. Mental effort is as truly labor as physical exertion. Probably the physician derives more pleasure from his work than a factory helper does, but our only standard of whether a given effort is labor or not is whether it assists in production. The efforts of the physician, the lawyer, the judge, the business executive, and the government official all add to the total amount of enjoyable goods and services which constitute the income of society. Some of it directly increases the flow of goods and services, some indirectly by making the work of others more productive.

We are all aware that the effectiveness of labor varies with different climates and with different races. The contrast between a backward country, located in the tropics, richly endowed by nature, and one of the Scandinavian countries or Iceland suggests itself. In spite of the relatively unimportant natural resources of the latter countries, the per capita addition of each man to the world's total production is much greater there than in many a more richly endowed tropical country.

Race and inheritance. One writer has suggested that the effectiveness of labor is dependent on three factors — race and inheritance, health and energy, training and social environment.[1]

---

[1] *Cf.* E. Huntington and F. E. Williams, *Business Geography*, p. 57.

That the qualities which any race inherits are of great importance cannot be denied. A high degree of intelligence, the ability to learn, the ability to persevere are characteristic of most northern races. Inability to receive more than a small amount of instruction or to grasp abstract ideas, fickleness of purpose, and a lack of the power of concentration are found in some of the races of Africa and Australia. Perhaps the development of the former characteristics has been the only means of racial preservation in the inhospitable northern regions, while maintaining life has been easier in the warmer climates. This speculation opens up some present-day problems of importance, but these cannot be considered here.

Health and energy. These seem largely to be governed by climatic conditions. Both a high average temperature and a low average temperature impede mental and physical efficiency. Investigation seems to show that physical ability is at its highest when the average out-of-door temperature is about 64° F.; that is, when the temperature rises to about 70° at noon and falls to about 55° at night; and that the optimum point for mental energy is an average temperature out-of-doors of about 40°; that is, a maximum temperature during the day of perhaps 50°, falling in the night into the 30's. Since the work of most men is a composite of the mental and physical it is necessary to combine these figures, and we find that the optimum average out-of-door temperature for both mental and physical work varies from 45 to 60 degrees. If the temperature rises above or falls below these figures a decrease in energy is noticeable. Other things being equal a variable climate is more desirable than a monotonous one. A rise or a fall in temperature, or a storm, seems to stimulate activity. From this standpoint as well as from the standpoint of temperature the climate of the northeastern part of the United States is highly stimulating.

Training and education. A good racial inheritance, health, and energy are of little avail without proper training. Such training is more important in the present era of the world's civilization than at any time in the past because of the greater complexity of the industrial and commercial processes. Even with the amount of training already provided there are doubtless thousands at work in positions for which they are not adequately trained and in which they are consequently not as effective as they should be. And perhaps more disturbing is the thought that many men and women are doing routine work in a highly efficient manner who have the capacity to handle work of a more exacting nature more satisfactorily than some of those to whom it is now entrusted. The work of the factory operatives as a whole is quite as essential to the productive processes as the work of the scientist. The reason we value the efforts of the scientist so much more highly than that of the former is because there are relatively fewer capable of becoming trained scientists; we value the work of the expert accountant more highly than that of the bookkeeper because anyone with a moderate amount of intelligence can learn to keep a set of books, while the ability to interpret the accounts and to devise systems which will accurately tell the history of a business concern is not given to all. The failure to make use of such talents as individual members of society have constitutes a social loss; production is not as efficient as it might be.

The most reasonable hope of improvement in the use of labor lies in an expansion of the field of training and education, and particularly in the field of vocational education. To find out what a boy or girl is best fitted by natural endowment to do, whether medicine or engineering or blacksmithing, and then to provide the proper training is at the same time one of the most difficult and one of the most important of modern educational problems.

Capital. The third factor in production is *capital.* Ordinarily the term capital is synonymous with wealth; that is, useful material articles owned by man, but in discussing the factors of production we limit the term capital to wealth which man has produced, thus excluding land. We distinguish land from capital in this classification because of its peculiar relationship to production, but it should be remembered that it constitutes a particular and important class of capital in the ordinary use of that term.

At this point we are particularly interested in those capital goods which are useful in production. There is an important group of capital articles which is desired primarily for its own sake rather than as instruments of production, including the mass of goods ready for consumption — food, clothing, furniture, etc. To this we give the name *consumers' capital,* while we call the tools, implements, machines, and materials which are to be used to produce consumable goods *intermediate goods* or *producers' capital.* These classes cannot be exactly distinguished, nor are they mutually exclusive. Some goods are capable of being used for further production or of yielding their satisfactions at once. Coal may be mined for use in factories, or it may be used in the kitchen stove with the double purpose of warming the room and cooking food, or it may be burned in the open grate. A ship may be capable of use as a yacht or as a carrier of freight. Moreover virtually every sort of consumers' capital may, when used to satisfy wants, be at the same time serving the ends of production. This is obviously true of the form of consumers' capital *par excellence,* food, and it will be recognized as true even of many of the less essential forms and even of many mere luxuries, such as athletic goods. But in spite of these qualifications we shall find the concepts useful in analysis, and with this warning the reader can easily avoid confusion.

Capital in production. We are so accustomed at the present time to tools and machinery of every description that it is difficult to imagine conditions in which they were entirely lacking; yet long before the beginnings of history man must have been dependent on the unaided efforts of his hands. Such a situation, where brute strength or cunning is relied upon to secure the means of subsistence, we call *direct production.* What the first implement was is a matter of conjecture; it may have been nothing more highly developed than a club. Then possibly someone conceived the idea of attaching a sharp stone to the club, and an axe or hatchet was the result. Certain it is that tools came into use very slowly, and for centuries only the rudest and most primitive were known. Even during most of the centuries of which we have historical records, the machines and tools were of the simplest sort. But with the first tool came a definite advance, for man had started on his yet unfinished process of changing the conditions of his environment to meet his rapidly growing needs.

Production with tools and implements is called *indirect* or *roundabout* production. One may for example acquire the art of catching fish with his hands — this is direct production; but if he take time to make a fishing net he can greatly increase the effectiveness of his labor, including, as of course we must, the labor spent in making the net. Instead of proceeding directly to secure the desired products, time and energy are first spent in making tools, with the result that the production is made vastly more effective and, passing beyond the small range of activities open to one without tools, things are produced which could not conceivably have been made with the unaided hands.

We are not however particularly interested in the contrast between direct and indirect methods of production; the economies of the latter method are so patent as to require no further demonstration. The contrast of significance is between a less and a more intensive resort to the roundabout methods of production. The rude plow which our ancestors used and which is still to be seen in some parts of the world was but a pointed piece of wood hauled by oxen and could easily be made by the farmer himself. Contrast the labor which must be spent in building the modern plow and tractor. First of all the ore must be mined, refined, and worked, the lumber must be milled, machines must be made for making the iron and steel parts, and machines for making these machines. There are in fact almost countless preparatory stages in the making of any machine at the present time. Yet even counting the labor of these preparatory steps the modern plow by far outdistances its progenitor in output. The labor necessary for the erection of a modern flour mill may be spread over many years, the materials for its construction may be gathered from the four corners of the earth, and hundreds of men may be engaged on one or another of the preliminary processes, whereas the old-fashioned gristmill could be made in a few weeks' time by a relatively small number of laborers; but the output of the modern mill per unit of labor necessary to erect and operate it is considerably greater than came from its predecessor. Wherever we find an example of intense recourse to the roundabout or capitalistic method of production, we may safely conclude that it is because the effectiveness of labor is thereby enhanced. Capital instruments are labor-saving devices, not only from an individual's point of view, but from the point of view of society as well.

The accumulation of capital. Saving. The degree to which indirect methods of production may be carried depends, in part at least, on the quantity of capital which is available for productive purposes. Other elements enter, such as the progress of invention, the size of the plants in the industry, the demand for handmade articles, and so on, but the quantity of capital is so fundamental a consideration that a study of the way in which it is accumulated will repay our consideration.

The formation of capital depends upon the readiness of men to postpone the gratification of wants in the present in the hope of increased pleasures in the future; in other words, it rests on *saving*. In relatively simple cases this is quite clear. The savage who has gathered a supply of corn may be tempted to consume it before the spring, but if he can save enough for seed the following year he can hope for a more abundant supply in the autumn. A farmer today may use the proceeds from the sale of his crops to buy either a new automobile or an agricultural machine. The former will yield him pleasure at once, but the latter will enable him to expand his farming operations and to have a larger surplus of grain in the future. Nearly every manufacturer is confronted at some time or other in his career with the choice between using the profits of the business to expand his scale of living and " putting them back in the business " in the form of enlarged facilities.

In these cases we can easily see that self-denial — saving — has added to the stock of capital. For the majority of people however the connection between saving and the productive processes is not so clear or so direct. Today the great mass of people gets its living by exchanging services for money and with this money it purchases the various articles which are required or desired. To the individual the saving of a part of his income does not seem directly to affect production; it appeals to him solely because of the necessity of providing against illness or old age or because under our present system money can be left in a savings bank or loaned or used to purchase stocks and bonds and can thereby earn interest. Yet it has precisely the same influence on production as saving by the manufacturer.

If we consider saving from the standpoint of society as a whole it will be apparent that it entails the devotion of effort to the production of intermediate goods which might have been spent in producing goods capable of yielding their satisfactions at once. It further implies that the consumption needs of those working on the intermediate goods must be met from the flow of consumable goods produced by others. Were it not so they could not live.

We can imagine a simple kind of social organization where the tasks are arbitrarily divided among the different men and women, where all share in the common product, where the farmers are required to set aside a portion of their produce to support the miners and others working on intermediate goods, receiving their recompense eventually in the form of agricultural implements which will greatly increase the effectiveness of their efforts.

But although we may be able to find isolated cases where the intermediate worker is supported by direct payment in commodities, such as "grubstaking" a prospector, it is far from the normal case. The ordinary individual saves money, not commodities, and the laborer receives his remuneration in the form of money. This feature serves to obscure the nature of the problem and make it seem more difficult than it really is. A clerk who saves $500 in the course of a year does so by denying himself the satisfactions which might have been obtained from the consumption of goods. When he places the sum in the savings bank, it is not kept in the bank's vaults until such time as he may require it but is loaned to some borrower. Now it is possible that the borrower may want the money to purchase goods for his own immediate pleasure, but the majority of loans are made to men actively engaged in business and for business purposes. Let us suppose that the savings bank uses the money to buy a bond, a part of an issue of bonds put out by a railroad in order to be able to finance the construction of a branch line. The railroad uses the funds thus secured to engage men to work on the construction of the line and to purchase the results of past labor — steel rails, ties, etc. The clerk has actually passed on his control over consumable goods to the laborers engaged in the intermediate processes; by reducing his own consumption he has made it possible for them to work on non-consumable goods and yet to live. During the years which may be required to construct the railroad line the men engaged in all the preparatory stages might have been producing food, clothing, etc., and for the time being have greatly increased the flow of consumable goods, but the eventual result of abstinence on the part of some in society is the creation of a great machine of production, the railroad, which makes the efforts of workers in all lines many times more fruitful. The saver gets his reward in the form of interest, but he also shares in the social gain, which consists of a more abundant flow of consumable goods.

We shall often refer to the importance of the capitalistic or indirect methods of production and the necessity of saving. It is difficult to overestimate the importance of either, and it is easy to overestimate the adequacy of the volume of savings at present available in the world. Were the people of the nations of the earth suddenly to stop saving, in the sense of spending all earnings on consumable goods, it would be but a short time before the entire structure of our civilization would be crumbling in the dust. On a smaller scale this has happened in Russia. In the days before the war Russia's industrial fabric depended in large measure on loans from foreign nations. Under the Revolutionary and Soviet régimes Russia has been cut off from foreign sources of supply, and in the disorder which has prevailed has been living in part on her capital. There has been practically no replacement of capital goods which have deteriorated or worn out. The result has been an enormous decrease in the efficiency of Russian labor and the threatened extinction of many lines of industry. The remedy is either saving by the Russian people and government or borrowing the savings of others.

The form of capital instruments. Invention. The effectiveness of labor in production depends not only upon the amount of capital available for use in production, but also upon the form which this capital takes. At the present stage in industrial development we could doubtless profitably use a greater number of the tools and machines with which we are already familiar, but there must be some limit to the gain which comes from duplicating our present machinery. A more profitable line of development lies in improving the type of machines. Not two printing presses of the same capacity, but one improved press with twice the ordinary capacity.

The inventive genius of the world is constantly studying the possibilities of improving old machines and devising new machines to take care of processes now performed by hand. Society profits by the inventions which make it possible to secure a greater flow of commodities with a smaller expenditure of effort. The saving which we have expressed in general terms appears to individual members of society in the guise of a lower price and therefore means that a given income represents command over a larger number of commodities than formerly. Measures which stimulate inventive genius are generally in the public interest. The inventor may make a fortune for a few years through the monopoly which a patent right confers on him, but the people as a whole benefit still more. In the long run the hope of making a fortune is the spur which leads the inventor to persevere through the period of fruitless experimentation. To those who understand, the wealth which Edison has accumulated appears a small price to pay for the services he has performed for society in placing so many comforts within the reach of people of modest means.

The capitalist class. Strictly speaking, anyone who owns capital is a capitalist, and there can be no clean-cut distinction between the popular terms " capitalist " and " laborer." Many a laboring man is the possessor of at least some small amount of capital, invested and yielding him an income in the form of interest or dividends. And the majority of the wealthy owners of great stores of capital are actively engaged in productive enterprise and so performing labor as that term has been defined. Yet vague though the distinction may be on the border line, it has become customary to speak of capitalists and laborers as separate social classes, and the classification and terminology have their purposes. In this sense the laboring class includes those who, though possibly owning some small portions of capital, nevertheless devote themselves to manual labor or the lower grades of mental labor and derive at least their principal support from the wages of their labor. The capitalist class includes those whose holdings of capital are relatively large and who are either above the necessity of performing labor or whose labor, if they do work, is of the higher grades devoted to direction and control of capital and the labor of others.

This class derives a considerable income in the interest and dividends from their capital, quite apart from any reward for their labor, and their position enables them to exert a powerful influence upon the destinies of the laboring class. To many the income and the power which result from the ownership of capital seem hard to justify. Justification or condemnation of the capitalistic system is not our purpose. But it is pertinent to call attention to the essential service which is rendered by the capitalists and which the careless critic sometimes forgets. Capitalistic or roundabout production is a time-consuming process. Someone has to wait until the final product, the good which has led to all this preliminary investment in machinery and building, emerges. During this interval the workers must live. They are in reality producing a future good, but they must have present means of securing present goods. The capitalist is the one who does the waiting, and his operations resolve themselves into a series of advances to the laborers. The logic of the return which the capitalist gets in the form of interest is based on the fact that without this incentive he will not wait, will not tide the laborers over the preliminary stages of production, but will increase his own consumption. Savings will be checked, and society will lose as production falls off.

Possibly some other system may some day be devised and carried into operation which will supplant the individual as the saving unit in society and hence do away with the necessity of a personal reward, but whatever system is attempted the waiting period can never be done away with. And whether socialism or communism or syndicalism be attempted, the capitalistic method of production cannot be destroyed without relapsing into barbarism. Either individuals in society or the group as a whole must save and turn the savings into the machinery of production.

The entrepreneur. Land, labor, and capital cannot work together effectively without organization, direction, and control. This function in the modern economic organization is entrusted to the business man, the enterpriser, the *entrepreneur*. So important is his function that the entrepreneur is by some writers elevated to the rank of a fourth "factor of production." Since whatever the entrepreneur does falls clearly within our definition of labor, there seems no need of thus putting him in a class by himself. He is economically a laborer, although in the popular separation of social classes he generally falls in the capitalist rather than the laboring class. In fact the entrepreneur is quite generally a considerable owner of capital and so a capitalist as well as a laborer. We are able nevertheless to distinguish between the functions of laborer, capitalist, and business leader, even though they may be united in one person.

The function of the entrepreneur is primarily that of making decisions as to the use of land, labor, and capital in any given business. He determines the relation of land, labor, and capital to each other, the degree to which it is profitable to supplant manual labor by labor-saving machinery, the amount of a given good which can profitably be produced. Whether a new invention is to be given a trial depends on his judgment. He studies the nature of the industrial operations, tries to simplify them as much as possible, and assigns laborers to one task or another according to his opinion of their abilities. The responsibility for the entire internal and external organization of a business falls on his shoulders — the coördination of production, finance, buying, and selling. The larger the business, the more incentive there is to delegate to others all tasks of a routine nature and the more closely does the work of the chief executive partake of the nature of decision-making and his reward consist of business profits.

Business profits tend to vary directly with administrative and executive ability. Many business men at the present time get no more than wages — the wages which would be paid any hired manager to attend to the routine of the business. Others get business profits which correspond to their superior business ability. There are those who regard the profits of the business man, particularly when they are large, with disfavor no less than that with which they contemplate the interest of the capitalist. Here again it is not our purpose either to justify or condemn the present system. But it is evident that much of the condemnation of business profits is the result of failure to appreciate the service which the business leader really performs.

As a matter of fact it is probable that as much native talent, not to say genius, is required of those who attain first rank in business leadership as is required of the pianist who becomes a performer of the highest order. There are many with modest musical gifts, but few with the natural ability, intellectual insight, and emotional qualities of Liszt or Rubinstein. We refuse to tolerate the public performances of indifferent performers, but we are compelled to permit many business concerns to be managed by men of mediocre business talent, because of the lack of a sufficient number of leaders of the highest grade.

It would seem therefore profitable to society to encourage in all ways the exercise of this business ability in order that the most capable leaders may serve as guides to the less capable. If the stage is ever reached where further progress is impossible, business profits will cease as the less capable of the business men finally catch up with the methods and devices of the more capable. As it is the more talented business man stays several laps ahead and maintains the disparity between his income and the incomes of others by trying out new and possibly risky methods when his old methods of organization and production have become the common property of all. Thus is secured progress in the development of more effective organization of the factors of production, from whose enhanced product the whole people stand to profit.

Wed Look up ~~[illegible]~~ concompetent
demesne
morain
coalesed

CHAPTER III

INDUSTRIAL STAGES

The industrial structure of today is so complex that it is difficult to comprehend it as a whole or to grasp the significance of its essential features. From various quarters critics are assailing the present order and demanding this reform or the abolition of that institution. Some of our zealous reformers would think it a small price to pay were destruction of the present social organization involved in bringing about the reforms they desire. Others, impressed with its notable services, proclaim the present order as little, if any at all, short of perfection. It is small wonder that in this din of verbal conflict the citizen who conscientiously seeks a fair appraisal of the economic system finds himself somewhat bewildered.

The purpose of Part I of this book is to enable the reader to examine the modern industrial structure in order to visualize and comprehend its most important features and to understand and appreciate the way the component parts of the organization work together to meet the needs of society. Only with such an understanding is one in position to pass judgment upon them.

As a foundation for this study it will be worth our while briefly to survey other systems which man has evolved in his attempts to solve the economic problems of life. This will serve two purposes. Not only will it enable us to understand the modern organization better by knowing something of its antecedents, but it will give a background which will be helpful in evaluating modern institutions and methods. The reader will be enabled to grasp the fact that material progress has come slowly and painfully; that the degree of prosperity to which we have already attained is not lightly to be tossed aside or thoughtlessly jeopardized. He will be likely to acquire a healthy skepticism toward those who would bring forth the millennium at once by the passing of laws, yet he will find abundant reason to anticipate future increase in material prosperity even more rapid than the past, if the factors which make for progress are not stifled by artificial restraints.

The household system. It will be sufficient for our purpose if we go back only to that stage when the family or household was the economic unit, ignoring the struggles of primitive man which lie back of that period. The family, using the word in its broadest sense to include the clan or the tribe, lived a life unto itself. Most of the goods consumed by the family were made by the members of the group, for there was practically no trade with outsiders. There must have been some division of tasks among the various individuals as natural aptitudes showed themselves, but the opportunity to develop special skill was very limited, and each individual was called upon to do a variety of tasks. Such a mode of life implies a limited number of commodities to meet the needs of life and a relatively low standard of living. Few families or households could be so situated as to command either the raw materials or the technical skill necessary for the fabrication of even a small fraction of the goods with which we are today familiar.

It was a simple type of organization indeed, lacking in many of the most common features of our modern life. There was virtually no exchange and hence little need for money; no wages were paid, as all were working for the common good and shared in the common product; the amount of capital used for production was limited to a few tools; the transportation involved was slight and gave rise to no problems such as vex the modern world. The organization was so foreign to our own experience that we have to turn to the remote past for an illustration. The large estates of the Romans might serve us as an example, but a more interesting example is at hand in the manor, which was common in Europe in the Middle Ages and was the political and economic unit in England in the eleventh century.

The manor. The manor itself consisted of a village surrounded by arable land, pasture, waste, and woodland. Over each manor was a lord, who held it directly of the king or perhaps indirectly of some overlord on the basis of certain services chiefly of a military character. Living in the village and cultivating the land which they held from the lord were tenants of various grades from the slave to the freeman.

The arable land was divided into two portions, that which the lord or his representative held, called the demesne, and the holdings of the tenants. The holding of a tenant might amount to thirty acres, while the holdings of the lord constituted a third or a half of the total arable land. Both the lord's and the tenants' holdings consisted of a number of acre and half-acre strips scattered over the tilled area, so that no one held his land in one block. The arable land was usually separated into three great open fields, one of which was planted to wheat or rye, another planted to barley, oats, or beans, and the third allowed to lie fallow. The following year the field that had been allowed to lie fallow would be planted to wheat, the wheat field planted to barley, and the third field allowed to lie fallow. This was the rather crude method of conserving the fertility of the soil, made necessary by the lack of fertilizer and of knowledge of the proper rotation of crops.

One class of tenants was known as villeins. They had somewhat the status of serfs in that they were bound to the soil and were compelled to cultivate the demesne of the lord and to perform additional services for him at other times of the year, notably at the time of harvest. But the villein cannot be considered a slave, for he had his holding of thirty acres with a proportional share in the use of the waste, pasture, and woodland; he could not be sold from his holding, he enjoyed an independent family life, and he had the right to property which he might acquire by his own labor.

The manor was a world in itself, with very little relation to the world without. While not a family in the blood sense, it was a self-reliant household with the lord at the head surrounded by a group of dependents. As can easily be imagined the most important occupation was agriculture, and it is no exaggeration to say that all of the laborers on the manor were agricultural laborers, although there was some slight development of crafts such as blacksmithing and milling. All the food for the group was raised on the spot, the clothing was manufactured from the wool of sheep grown there, and the tools necessary for the cultivation of the soil and for such primitive arts as existed were of necessity the product of local effort. Trade in the modern sense scarcely existed, save in a few most necessary articles. Salt was always required, particularly to preserve meat through the winter, and in most cases it had to be procured from without. Tar was purchased to keep the murrain from the sheep. Animals also could be traded in since they provided their own means of transportation, but in general the difficulties of moving goods about were prohibitive. Even the kings and rich nobles who were the holders of several manors were compelled to come to the manors and consume the produce on the spot, since it could not be conveyed to them. The wage system was not to be found, nor can the payments which the tenants made to the lord be considered the equivalent of modern rent, as they rested on custom and not on any estimate of the worth of the land to the tenant.

Life on the manor may be made to seem attractive if sufficiently idealized, but as a matter of fact the modern worker would shrink with horror at the thought of being condemned to such a life. The semi-servile worker, who typified the largest proportion of the population, could not leave the spot where he was born. His food was coarse and monotonous, his clothing the product of rude home industry, and his shelter squalid. Comforts, refinements, and pleasures simply did not exist for him. His wants were doubtless few, and fortunately so, for the means of gratifying wants outside those essential to life were entirely lacking. Such contentment as he found was that of an unthinking animal which takes what it finds and looks no further.

The Handicraft Stage. With the growth in number and size of the towns we enter upon a new stage of development which has been called the Handicraft Stage. In the preceding period towns of importance were very few in number, but in the fourteenth and fifteenth centuries both on the Continent and in England they multiplied in number and in power. Many of them were founded under the protection of some powerful lord or ecclesiastical establishment, others seem to have grown up where trade routes met. In the early days of the town the population was primarily agricultural, the citizens cultivating their fields outside the town walls, and throughout medieval times the towns maintained in part their agricultural characteristics. But with the growth of the towns in population there came opportunity for certain men to develop special skill in particular crafts or trades, and we find artisans devoting the major part of their time to shoemaking, others to weaving, and so on. The products might be sold to townspeople, to people from the countryside, or even to traders from other towns. For the town as a whole the development of craft specialization meant the exchange of the surplus of its manufactured goods for the surplus produce of agricultural districts or of other towns.

Merchant gilds. That both the merchant and the artisan occupied places of greater significance than formerly can be seen in the formation, first by the merchants and later by the members of the different trades, of associations or societies, called *gilds*. How far back the merchant gilds go is uncertain, but from the eleventh century on, as trade began to grow in volume and importance, there was a growth in the number and influence of these associations of merchants. Originally they were voluntary associations, entered into very likely for the purpose of mutual protection at a time when law and order had not developed to the point where it was safe for a merchant to travel about the country alone. Later on however we find them monopolizing the trade of the town and associating themselves with municipal authority in the attempt to maintain the monopoly of trade against outsiders. Indeed the relationship in many cases became so close that the merchant gild and the town authority coalesced.

Craft gilds. Shortly after the rise of the merchant gilds, craft gilds developed rapidly, and it was soon true that every craft or trade requiring any skill gave rise to a gild. In all probability the origin of the craft gild is to be found in a community of interest in religious and social matters; the members of a craft lived in the same district in the town and attended the same church, and it is natural to find them uniting perhaps to aid the widow of a fellow craftsman or to celebrate some religious festival. Probably the example of the merchant gilds was not unimportant in their formation. Whatever their origin, however, we soon find them devoted primarily to the regulation of matters pertaining to the craft, although the social features were always retained.

In theory at least each gild regulated its own affairs in the public interest, and to ensure greater effectiveness in this a monopoly in each craft was granted, so that a man not a member of the gild might not ply his trade in the town. This requirement was not burdensome at first, for the gilds were interested in acquiring greater influence through a large membership, but later, as the position of the gilds became more firmly established, the monopoly power was used to protect the members against the possible competition of others, and admittance to the gild was made very difficult. Chief among its regulations in the interest of the community were those intended to secure honest workmanship. The right of gild officials to inspect the work and the workers' premises was recognized, and fines were levied for infringement of the gild regulations and for practices intended to deceive the purchaser as to the quality of the ware. Partially to secure youths thoroughly trained in the mysteries of the craft and partially to limit the number of men in the gild, a system of apprenticeship was instituted, so that a youth who desired to become a master weaver, for example, was required to serve with a master for a period, ordinarily of seven years. At the end of that time he might be admitted to the gild and allowed to set up his own establishment, or he might continue to work for his master as a journeyman.

The craft gild presents a superficial resemblance to the modern trade union, but essentially it was quite different both in its constitution and in its aims. It is true that the aim of the gild was in part directed toward advancing the interest of its members and that the methods pursued closely resemble those to be found in some of the most highly organized trade unions of today, but as we have seen, the primary aim of the craft gild was to ensure honest workmanship and honest goods. Furthermore the trade union is composed of men working for an employer on machines and materials furnished by him and receiving wages, whereas the membership of the gild was made up of independent artisans, working at their own pleasure with their own tools and materials and securing their remuneration by the sale of their finished goods.

Municipal regulation of commerce. In this period the town rather than the manor was the economic unit. On some parts of the Continent the town tended to be the political unit as well as the economic unit, but in England, in theory at least, the authority of the king was superior to municipal authority. In practice the towns enjoyed a degree of local autonomy which is inconceivable to us. The authority of the town and the sympathy of its citizens were all on the side of increasing the wealth and prestige of the town. It was a patriotic duty to protect it against the possible competition of the countryside and of other towns. Only if we remember this fact can we understand the spirit animating municipal legislation in this age.

Among the most important of such regulations are those governing commerce. Commerce was conceived of as a privilege of a citizen of the town, to be extended to others only grudgingly and with reservations. Each town had its customs barriers which prevailed against other towns and the countryside. A foreigner (who might be from the next town) might sell his goods at the market but only under surveillance. He must sell only to townsmen, or to others after the citizens had satisfied their requirements; he must then have a citizen of the town act as intermediary, lest two foreigners conspire against the interests of the town, and he must expose all of his goods at once so that the total stock might be known. There were many other regulations of a more general nature to protect the interests of the townspeople. Selling goods on the way to market was forbidden; buying up the entire supply of a good to sell at a higher price and buying to sell again without adding anything to the value were likewise forbidden. The prices at which certain goods might be sold, the manner in which debts might be collected, the system of weights and measures, and a host of other matters directly or indirectly connected with the town trade were controlled by town authority.

The beginning of specialization. If we contrast this period with the one preceding, it is evident that there have been changes of considerable moment. In the first place there is the beginning of specialization in production, of the differentiation of one occupation from another. Individually each is less self-reliant; he depends in part at least on the productivity of others, and secures some of the articles he uses by making more of one article than he can use and exchanging his surplus for the goods of others. This development of exchange is significant, for exchange involves in some manner an evaluation of the two products, an agreement as to the relative worth of them, in other words value. The market, a place where buyers and sellers can meet to determine price by exchanging goods, was also created. And a still more fundamental change can be seen in the growing necessity of money as a medium of exchange. Money had been used previously for the payment of fines and for hoarding, but with regularly established habits of exchanging goods it became increasingly desirable. If one purchases few goods he can doubtless exchange wheat for shoes and cloth after a fashion, but if he depends on exchange to secure his whole stock of goods there must be some common unit in which the prices of all goods can be expressed.

Growth of freedom. It should not be thought that the development of town life was without its influence on life in the country. The citizens of the town were free as opposed to the semi-servile status of the rural inhabitants. Before the growth of the towns there had been little chance for an agricultural laborer to improve his condition, for there was no place where conditions were better. If he ran away to another manor he would either be returned to his lord or might have to accept even worse conditions than those from which he had fled. Now he could go to a town, and, as the custom held, if he remained there a year and a day he became a freeman. This raised the status of the agricultural laborer, as it made him more valuable in the eyes of his lord. Increasing use of money and familiarity with money payments, together with the demand from the towns for the surplus which the tenants might produce, brought about another change. This was the substitution of money payments for the personal service on the lord's demesne.

There were other factors and forces working in the direction of raising the status of the agricultural workmen. Into these we cannot enter, but we must note the result. In some Continental countries serfdom lasted until the eighteenth and nineteenth centuries; in England it had died out by the close of the sixteenth century. This is of significance, not merely on sentimental grounds, but because a free population is more energetic and ambitious than a servile one. England's development in the seventeenth, eighteenth, and nineteenth centuries was in large measure predicated on the superiority of a free population.

Judgment of the town economy. It is safe to say that the period of town economy marks an advance over the previous stage. Certainly in the realm of material things there had been progress. The specialization of workers and the growth of trade brought many articles of commerce within the reach of the country population which they had not known before. The influence of the towns in raising the status of the agricultural laborer was decidedly progressive. The towns themselves offered means of amusement and enjoyment which had been lacking entirely or open to relatively few. On the important religious festivals the gilds united in giving "misterie" plays, so called from the name "misterie" or craft (from the French *métier*), to the delight of themselves and their audiences. The market was a source of profit not only in providing a place for the exchange of goods, but in bringing people together, and some of the towns had periodical fairs, which attracted merchants and consumers from a wide area. There was therefore manifest an increased productive capacity and the possibility of satisfying a wider range of wants, and, not least of all, a wider mental horizon.

It would be a mistake however to think of this epoch as the golden age, as some are wont to do. A medieval town was at all times picturesque, but probably at no time an attractive place of residence. The streets were excessively narrow, unpaved in many cases, and served not only as thoroughfares but as the dumping ground for refuse of all sorts. The houses were damp, cold, badly lighted and ventilated, and in many cases indescribably filthy. Plagues were a natural and inevitable concomitant of the unsanitary arrangements. Perhaps the independent craftsman got more pleasure from his work than the factory hand of today, but even that can be questioned. Certainly we know nothing which would invite us to try to return to the handicraft system.

The domestic system. As time went on there was a development along two lines. In the first place a greater degree of specialization by the individual in his particular trade manifested itself, and secondly some of the crafts were divided into distinct trades. This can be seen in weaving, where fulling and dyeing became separated from weaving. These changes were doubtless imperceptible at the time, but by the sixteenth and seventeenth centuries we see industry organized on a new basis.

Relatively little change is noticeable in the technique or methods of manufacturing. In weaving the workers still work in their homes at their hand looms with relatively few assistants. We see however that the craft gilds have lost some of their power over the trade and that weaving and other crafts have found roots in the countryside. But the chief change has come in the control and direction of industry. The worker is no longer an independent craftsman working directly for the consumer but executes orders given him by a merchant, who supplies the materials on which he works and in some cases the looms as well. The craftsman has become a wage earner.

The explanation of this change is somewhat obscure in its details, but the general outline is clear. In the handicraft stage the artisan worked for a market which was at hand. The consumer himself gave him the order for the cloth or the boots. In the course of time the market for his wares expanded with the decreased isolation of the towns and with the increased importance of English foreign trade. Instead of production for a known and familiar market, production for an unknown and distant market was being substituted to an increasing degree. The craftsman was a maker of goods, not a merchant. He had neither the knowledge nor the ability to handle the marketing of his wares in distant parts, and there was therefore a definite call for the services of a middleman. The merchants stepped in and shouldered the risks of marketing.

Foreign trade was expanding, large profits were being made by the merchants, but the opportunities for investing these savings were rather limited — confined in the main to lines with which they were personally allied. It is but a natural step from buying the products of a manufacturer to giving him orders, supplying the material, and even furnishing the instruments of production. Further profits could be made by stimulating production, and savings were available for this purpose. Some even went so far as to collect workers in separate buildings, thus anticipating the factory system, but in general there was only the modification of the handicraft system which has been described.

By the early part of the eighteenth century this system was common in all lines of industry, although the independent small master survived in some lines, notably in certain of the woollen districts, and it lasted until it was supplanted by the factory system in the early part of the nineteenth century. Examples of it survived in the United States down to recent times, particularly in some phases of the garment industry. Vests and pants were given out, chiefly to newly arrived immigrants, to be stitched or buttonholed in the home at very low rates of pay. The exploitation of the immigrant in this way became at one time notorious. On the Continent in some of the specialized branches of the textile industry, such as lace making, it still persists.

This period is of peculiar significance, not only because of the breakdown of the customary methods of production, but because it is the beginning of capitalistic production in a new sense. We have already defined what is meant by capitalistic production, and of course production in the handicraft period was capitalistic. The difference here lies chiefly in the divorce of the capitalist and the worker. The master craftsman worked with his apprentices and journeymen. The merchant-manufacturer, as we may call the member of this new class, did none of the manufacturing himself, but he owned the materials and the finished goods and assumed all the risks in connection with production. It is the beginning of the separation of the worker from ownership in the tools of his craft, a step which is definitely taken in the factory period, and one which is responsible for some of the bitterness of modern labor problems.

The factory system. The factory system, which is characteristic of present-day industry, came to the fore in the last part of the eighteenth and the first part of the nineteenth centuries, though isolated examples of factories are known much earlier. Its advent is closely connected with that movement or development to which the name of " Industrial Revolution " has been attached. We shall study the Industrial Revolution in the following chapter and naturally shall have to consider the causes of the growth of the factory system and its chief characteristics. For the sake of completeness however we may here briefly contrast it with the domestic system and point out its main features.

In the first place there is the development of machinery. It must not be supposed that in the period of the domestic system there were no machines; the spinning wheel, which is as truly a machine as the most complicated modern automatic spinning machine, was known and used at a very early period in history. But now the simple machines worked by hand, which were all that existed in the previous stage, are replaced by complicated power-driven machinery. In the second place the machines, instead of being in part at least the property of the workers, belong to the factory. Thirdly in place of working in the home the workers are gathered together in a factory building. This implies a regularity of hours of work and a discipline during work which did not exist previously. And in the fourth place the factory system brings with it a still greater separation of tasks into minute operations, many of which can be performed by machinery.

The factory system has given rise to many of the problems of modern economic life. The wage system, which is an inevitable part of it, the divorce of the worker from the direction of his work, the decline in value of the technical skill of the worker as the consequence of the increased use of machinery — all of these and other features have created problems for which no satisfactory solution has yet been found. Yet as we look over the period covered by our survey, more than a thousand years, we cannot escape the conclusion that progress has been steady, that man has improved his lot. Increased production both in quantity and in variety has given him a greater range of satisfactions; steady progress from a semi-servile status to that of a free citizen has been accompanied by an enlargement of the scope of his activities and a widening of his range of interests. Too narrow a view of the strife and struggle of conflicting forces of the day may convince one that the world is going to the dogs. A broader view gives greater hope for the future, for man seems to have that within him which keeps him continually striving for better things, material, intellectual, and spiritual.

What con
center of population of your state?
Employer & Employee relationships
Competition for markets
government regulation of industry
Location of industrys. name all the

CHAPTER IV

THE INDUSTRIAL REVOLUTION

It will be remembered from our previous discussion that the domestic system of manufacture was current in England during the first part of the eighteenth century, a system which involved scattered workers, using their own or hired machines in the cottages in which they lived and working on materials supplied by a capitalistic manufacturer.

In the course of the eighteenth century there were changes in the technique of production of so profound a nature as to revolutionize industry. The period in which these changes came is called the Industrial Revolution. The term "revolution" is used, not to indicate a sudden change, but to denote the character of the change, for the transition, which started around the year 1770, had not completely altered the character of industry until about 1840 or even later. The Industrial Revolution involved the displacement of the domestic system of manufacture by the factory system, a great extension of division of labor, the invention and perfection of machinery in all lines of manufacturing, mining, and transportation, and consequently an enormous increase in the output of industry. It further involved social changes of great consequence, the congregation of workers in large factory towns, the beginnings of organized labor, the rise of a new class of capitalists and in England a new political alignment. In almost every phase of social, political, and economic life conditions were altered, new relationships brought into being. Most of the problems of economic and social life of today find their origin in this period, and the student who would understand modern problems and try to evaluate modern institutions and relationships must begin his study by a survey of conditions in England in the eighteenth century.

Why the Industrial Revolution began in England. Let us see first of all why the Industrial Revolution came in England and not in some other country, why conditions in England made her particularly ripe for a change. In the first place, contrasting England with the countries of the Continent, there was personal freedom. Serfdom had disappeared in England by the end of the sixteenth century, while it persisted on the Continent through the eighteenth century. In the rural parts of the Continental countries it would have been difficult if not impossible to secure the necessary labor force for work in the factories, and the use of water power made location in the country necessary at times. The gild system still had control of industry in parts of the Continent, and the gilds bitterly resented any innovations which might threaten their supremacy, whereas in England their monopoly of the trades had pretty well broken down by the seventeenth century. This made it possible for the business man to embark on enterprises as he saw fit, without being controlled and disciplined by members of his own craft. The relative freedom of the English producer from government regulation and restriction, as compared with the Continental producer, favored experimentation with new modes and methods of production.

The spirit of freedom was also manifest in the abolition of customs barriers between towns and in the political and religious security of the individual. England had made not a little progress industrially by serving as a haven of refuge for skilled workers suffering from religious persecution on the Continent. Finally her position of geographical isolation rendered her free from fear of foreign aggression and contributed toward making her a favorable field for industrial changes.

There were still other factors of equal importance. England had a large accumulation of capital, secured in the main from her foreign trade, which could be invested in manufacturing. Factory production with the use of expensive machinery is necessarily roundabout or capitalistic production. The foreign markets which the British merchants were serving were large and expanding; this meant that if more goods could be produced they could be sold. There were men trained in trade on a large scale, and production on a larger scale than formerly would not offer the difficulties that it might elsewhere. Finally there were vast deposits of iron and coal located conveniently to each other. Probably the absence of any one of these factors would have retarded the development which took place, but coal and iron were without question essential, for they were fundamental to the application of power machinery to mining, transportation, and manufacturing.

Coal and iron. During the early part of the eighteenth century the iron mining industry of England was in a period of depression, largely because of the difficulty of securing charcoal for smelting the ore. In all England there were in 1737 but fifty-nine furnaces, which produced 17,350 tons of pig iron. In the same year about 20,000 tons were imported, chiefly from Sweden. A cheaper and more efficient fuel had to be discovered before any great expansion in the production of iron could be expected. Coal had been known for some time and was used to a slight extent for domestic purposes, but it had been unsuccessful as industrial fuel. The urgency of the demand for fuel led however to further experiments, and, with the development of a method of coking the coal and the invention of the blast furnace, coal assumed an importance in industrial life which it has maintained to the present time.

The practical solution of the fuel question led to another problem, how to get the coal out of the mines. So long as little coal was mined, shafts near the surface could be used, but the sinking of deep shafts was impracticable because of the flooding with water. This difficulty was partially solved by the invention of the steam engine early in the eighteenth century and the improved steam engine of Watt later on, which made it possible to pump out the water as well as to haul up the coal by power. The development of coal and iron production in England is shown in the following table.

Industrial revolution

TONS OF COAL AND IRON PRODUCED IN ENGLAND[1]

Coal		*Iron*	
Year	*Product*	*Year*	*Product*
1700	2,148,000		
1750	4,773,828	1740	17,350
1770	6,205,828	1788	68,300
1790	7,618,728		
1795	10,080,300	1796	125,079
1854	64,700,000	1854	3,100,000

The figures for coal and iron are not given for the same years, but there is a fair basis of comparison. In the last half of the eighteenth century the production of iron increased seven-fold, that of coal more than doubled. In the first half of the nineteenth century iron production increased twenty-four-fold, that of coal somewhat over six-fold. These figures are interesting not only in showing the absolute increases in both iron and coal, but also in pointing out the increased efficiency in the use of coal. When we realize that coal was being used for power to an ever increasing degree, a twenty-four-fold increase in iron production coupled with a six-fold increase in coal production points clearly to improved processes.

The textiles. England's greatness in foreign trade and her domestic prosperity in the sixteenth and seventeenth centuries have been ascribed by many contemporary writers to her woollen manufactures, and it was a cardinal feature of her commercial policy to protect her domestic wool supply, her manufactures of wool, and her foreign markets. A part of the woollen manufacture was domestic in origin, but it had been greatly stimulated by the influx of refugees from the Continent. The immigration of Flemish weavers who settled in and around Norwich in the sixteenth century had been encouraged and had helped develop that region into a prominent manufacturing centre.

The cotton trade in the first half of the eighteenth century was insignificant and was confined to Lancashire, later to become the seat of an enormous and profitable cotton industry. Most of the raw cotton came either from the Mediterranean and Oriental countries or from the West Indies, but the amount imported was small. Yet the first important inventions in textile manufacturing were applied to cotton first and then slowly to wool. The reasons for this are not difficult to understand. The cotton industry was new, and there was maladjustment between spinning and weaving. For the former the supply of laborers was insufficient, particularly with the growing demand for domestic cottons consequent on the limitations placed on the import of India cotton goods. There was a famine in cotton yarns; the weavers could easily use more yarn than could be turned out by the primitive spinning methods. The development of cotton planting in the United States and the assurance which the invention in 1793 of the cotton gin gave of a continuous supply of raw material stimulated investment in cotton manufacturing. Furthermore the fact that it was a new industry gave the producers more freedom, for they did not have to face the hostility which would unquestionably have arisen in the case of the older and better established industry. It was probably easier to apply the new inventions and methods in the woollen industry after they had proved successful in the cotton industry.

---

[1] Knowles, L. C. A., *Industrial and Commercial Revolutions in Great Britain during the Nineteenth Century*, pp. 70–71.

Spinning. Spinning presented the greatest immediate problem. The old-fashioned spinning wheel which drew and twisted the fibres into a thread could make only one thread at a time, and it is said that six spinners were required to supply yarn for one weaver. It is therefore natural to find the first inventions in current use in this branch of the industry. In 1767 Hargreaves invented a hand machine, the spinning jenny, so named in honor of his wife, which enabled a spinner to make eleven threads at once instead of one and which was soon improved so as to be able to spin one hundred threads at once. This was followed the next year by Arkwright's water frame, a power machine as the name suggests, which spun a stout cotton thread that could be used for the warp (the thread which is placed lengthwise in the loom and which is crossed by the woof) in place of the linen thread previously used. And in 1775 Crompton perfected his mule, a hand machine, though soon adapted to power, which was a combination of the principles of the jenny and the water frame. This could spin thread fine enough for muslins.

Weaving. The increase in the output of yarn made possible by the improved spinning machinery and the increase in the supply of raw cotton reversed the situation which had existed earlier. Instead of a famine in yarns there was a surplus great enough to admit of a considerable volume of exports — a thing considered highly undesirable as it tended to stimulate foreign weaving. Attempts were made to attract weavers from the woollen trade, but attention was also focused on improving the weaving machinery. Weaving offers technical difficulties. As has been explained, the woof is passed through the warp, and to do this the warp must be capable of adjustment so that the woof passes over one thread and under the next; the succeeding thread of the woof will reverse the process, thus binding the whole together. In 1733 Kay had invented his flying shuttle, which simplified the task of passing the woof through the warp and enabled a man to do his work in half the time and one man instead of two to weave the widest cloth. It met however with opposition and was not introduced until the latter part of the eighteenth century. Cartwright had invented a power loom in 1784, but it was imperfect, and it was not until the first part of the nineteenth century that a really successful power loom was devised.

The use of power and the growth of factories. Factories with power machinery were not an immediate result of these inventions. The first of the new devices were hand machines and were used in cottages, as had been the case with the earlier hand machines. It is stated that spinning continued to be carried on in the homes on the jennies twenty to thirty years after they were invented. But as the spinning machinery was improved and became heavier, power was substituted; first of all horses were used, then water power, and then steam.

The use of power almost inevitably meant the factory system, with machinery, power, and materials supplied by the owner, and the laborers under a common discipline. In the first place the machinery was becoming too expensive for the workers to purchase; it had to be supplied by one with large capital resources. Again to use power economically it had to be applied to all machines at the same time, and this involved restraint on individual freedom, the fixing of common hours and conditions of work.

The factory system spread very slowly however. The new machines did not always work well, and when water was the source of power it offered serious drawbacks. There might be an insufficient supply, an oversupply, or no supply at all. It is said that by 1790 there were only 150 water frames in use, so great was the disinclination to start factories.

The improvements in the steam engine effected by Watt in 1776 and 1782 offered an improved source of power, and steam gradually supplanted water. But the steam engine was by no means the steam engine of today; it was imperfectly built, broke down frequently, and was not especially economical in the use of fuel. In 1800 there were steam engines in only fifty-three manufacturing centres, and as late as 1835 there were 1,297 water wheels to 1,953 steam engines.

It has been stated that woollen manufacturing did not apply the new machinery until after it had been used in cotton manufacturing. Several reasons for this have already been suggested, but there were other reasons as well. Wool offered greater technical difficulties than cotton, in that its fibre is more difficult to handle. Woollen fibres are softer, less uniform, and less tractable than cotton fibres. There are great differences in wools, and every fibre has individual peculiarities to a greater extent than is true of cotton.

There was also not the scarcity of hands which was manifest in the cotton trade. Throughout the centuries in the growth of this trade a fairly accurate adjustment had been made between the raw material and the workers necessary to make it up. The increase in the supply of raw material due to the imports of Australian wool did not come until after 1830. Nevertheless the use of improved machinery and the development of the factory system were inevitable. The jenny for spinning woollen yarn came into use about 1785, and power spinning in factories in worsteds had started by the late years of the eighteenth century, although it did not develop in wool spinning until early in the nineteenth century.

In speaking of the development of the factory system in the woollen and worsted industry one writer says: " . . . the adoption of factory organization and the introduction of machinery came very slowly. There were scarcely twenty factories in Yorkshire in 1800; the power-loom was not introduced into Bradford till 1826, when it was the cause of fierce strife and riots; combing was done by hand until well into the forties, and many technical difficulties rendered it undesirable to use the power-loom in the woollen industry until about 1850. Writers in the middle of the last century speak of the wide-spread existence of the cottage system, and the memories of people still alive reach back to the days when the hand-loom was to be found in almost every cottage. Thus we come to the conclusion that the Industrial Revolution had little more than its beginnings in the eighteenth century. The great change came first in the cotton industry, then in the manufacture of worsteds, and lastly in the making of woollen cloths. In the Yorkshire branches of the textile industry, the revolution did not actually take place until the nineteenth century; the face of Yorkshire had been little altered by 1800, and a half a century had still to elapse before it could be claimed that the factory and power-driven machinery had displaced the old hand methods."[1]

It must not therefore be thought that the growth of factories meant the entire elimination of the home spinner or weaver; they persisted well into the nineteenth century. As late as 1833 a parliamentary committee estimated that there were 200,000 hand-loom weavers of cotton in the United Kingdom. In the beginning the factories would seem to have supplied the increase in the demand, leaving most of the weavers unmolested. Eventually of course the superior economies of factory production made it impossible for a hand weaver to get a living, but there were many side lines in which hand weaving persisted in to the second half of the nineteenth century. The stocking industry for example had not been transformed into a factory industry, and in the forties it was doubted if it ever would be, so great was the difficulty of applying power machinery to the making of stockings, and so firmly attached to their present mode of living and so opposed to going into factories were the workers.

---

[1] Heaton, *Yorkshire Woollen and Worsted Industries*, p. 283.

The reluctance to establish factories. The slowness in the establishment of factories may seem strange to those of the present. The economies of the factory system of production seem so obvious that it is difficult to appreciate the real opposition which existed a century ago. Let us remember in the first place that we are accustomed to change, that we welcome changes which have an even chance of proving profitable. The worker of the eighteenth century had no such viewpoint. The customary method of doing a thing was the right method, and violent changes were to be avoided. Furthermore the worker was fairly well satisfied with the situation as it existed. He might have a small plot of ground with his cottage and do a little farming on the side; he possibly had the right to pasture his cow in the common pasture. His wife and children worked with him and were self-supporting, and the family wage was fairly large and stable. He was his own master, could work when and how he chose, and found the discipline of a factory very distasteful. It is small wonder then that the factory was looked upon with disfavor as a disrupting force and that riots to destroy the factory machinery broke out in several places.

The opposition to factories appeared however not only on the part of the workers. The employers were none too eager to try them out. The "putter-out" of goods was really in a strategic position. He incurred no loss if he stopped giving out orders. His capital outlay was small compared to that necessary in a factory, unless he owned the machines. In that case he secured a rental from the worker and could at times use his position as a giver of orders to force the worker to pay a high rental. He was only gradually forced to turn to the factory system, either by a shortage of labor in certain lines or because of dishonesty in the use of the material which he supplied or difficulty in securing the goods when he wanted them. In some cases strikes on the part of the domestic workers caused him to establish a factory. But when the process was once under way the success of the great leaders forced others, employers and employees, into the system.

The development of the factory system, granted the inventions and improvements necessary to make it possible, was not due to the enthusiasm of the age; it was rather the result of economic factors which were well-nigh irresistible in their force. And of these the chief factors can safely be said to be the increasing foreign demand for goods and the necessity of supplying these markets in order to provide means for continuing the continental wars.

Developments in other lines. We have thus far treated the Industrial Revolution as if it were confined to the textile industries. True the development of the textile industries and particularly of the cotton industry offers an impressive example, but many other industries as well were affected by the Industrial Revolution. Indeed the development of the textile industries depended in a very vital way on the developments in other industries. Throughout the whole period there was constant action and reaction. Advances in one line caused advances in another line; these advances made possible still further progress in the first.

The early hand machines, the looms and the spinning wheels, were made of wood, and the first of the new spinning machines were also made of wood. The disadvantages of wood are in the first place the size which is necessary to secure strength, the impossibility of permanent precision in moving parts, and the difficulty of securing the requisite strength whatever the size. The latter difficulty was especially prominent when water and later steam power began to be used. Wooden machines could not stand the strain, and therefore first the movable parts and later all the parts of the machines were made of iron.

But the fabrication of iron machines, either for the textile industry or in the form of steam engines, was a new industry. There were few trained workers, the tools for making machinery were inadequate, and it is not surprising that parts made by hand did not fit when put together, or that continual breakdowns were the rule. It is said that when a new machine was sent to a factory the foreman of the shop which built it might have to spend a month fitting the parts together, easing up here and filing off there, until all the parts of the machine were adjusted to each other.

It was essential that machine tools should be made before any great development in the building of machines could be anticipated, for without machine tools precision in cutting and boring, making parts which are interchangeable, is impossible. The development along this line did not come until the first part of the nineteenth century. In part the demand for engines and machinery stimulated this development, in part was stimulated by it.

Transportation. The growth of the engineering industry and the increase in the output of goods of all sorts put a strain on the transportation system in two ways. There was in the first place the question of transporting the finished goods and the raw materials, and in the second place the hauling of coal from the mines and to the centres of manufacture.

The roads of England in the early eighteenth century were in very bad condition, and the naturally slow method of transport by wagon was rendered still slower. There were, it is true, waterways on which goods could be shipped, but they were not sufficiently developed nor well enough linked up with each other to suffice for the growing needs of commerce and industry.

The first step in the development of transportation was improvement of the roads. In the eighteenth century turnpike trusts were created which gave to private persons the right to charge tolls for the privilege of passing over the particular stretch of road. They had the corresponding duty of building the road and keeping it in repair. The roads which formed the main highways were kept in fair shape, though naturally there was a great deal of difference between roads, and in winter most roads fell into a shocking state of disrepair. At the opening of the nineteenth century road-making received a great impetus from the improvements of Macadam, who showed how a durable road surface could be made, and from Telford, who was skilled in road engineering. In addition the turnpike trusts tended to form amalgamations and to reorganize so as to operate more efficiently. During the first part of the nineteenth century there was steady progress in road-making and repair until the advent of the railway plunged many of these trusts into bankruptcy.

Roads can satisfactorily serve as a means of transportation for manufactured goods which have a fairly high value, but they are prohibitively expensive for bulky raw materials. At this time there was a growing demand for coal for industrial and household use, the potteries required clay, the factories in Lancashire needed a cheap method of transporting raw cotton. This need was met by the building of a network of canals, which, together with the rivers, joined the coal-producing areas, the industrial areas, and the sea. The first canal was built in 1761 by the Duke of Bridgewater to connect his coal fields with Manchester and was soon followed by another canal connecting Manchester and Liverpool. Then came a rapid succession of canals, built by private persons who collected tolls from the users. By 1830 there had been built 1,927 miles of canals and 1,312 miles of improved riverways, and there were also 812 miles of open rivers in England and Wales. A barge of twenty tons could navigate from one end of England to another. The result was a cheaper and more rapid method of transportation. Without the canals the development of industry would probably have been considerably delayed.

The railway period, which really opened in 1826, had been preceded by experimentation in rail wagon carriers using either gravity or horse power. The rails were constructed of wood until 1767, when iron rails began to be substituted and proved to be much superior. There had also been experimentation with stationary engines to haul cars. These early railways were built to carry coal from the coal mines to the canals or in some cases to towns. The Stockton & Darlington Railway was chartered in 1821 and was the first railway to use steam locomotives and to carry passengers, but the Liverpool & Manchester Railway, which was chartered in 1826, inaugurated the steam railway period. In 1829 Stephenson's "Rocket," though not the first locomotive, was the first to demonstrate clearly the practicability of the locomotive.

The development of the railway transportation system in Great Britain is a subject into which we cannot enter, although it forms a very interesting and important chapter in railway history. It is worth noting however that the railway was introduced and extended in England under great difficulties and in the face of much opposition. Some of the current objections to the steam railways are given in the following paragraph, and, although these arguments seem fanciful, it should be realized that they had great weight at the time and served to increase the difficulties which the promoters of railways had in securing charters from Parliament.

" 'The country gentleman was told that the smoke would kill the birds as they passed over the locomotive. The public was informed that the weight of the engine would prevent its moving; and the manufacturer was told that the sparks from its chimney would burn his goods. The passenger was frightened by the assertion that life and limb would be endangered. Elderly gentlemen were tortured with the notion that they would be run over. Ladies were alarmed at the thought that their horses would take fright. Foxes and pheasants were to cease in the neighborhood of a railway. The race of horses was to be extinguished. Farmers were possessed with the idea that oats and hay would be no marketable produce; cattle would start and throw their riders; cows, even, it was said, would cease to yield their milk in the neighborhood of one of these infernal machines.

" 'Vegetation, it was prophesied, would cease wherever the locomotive passed. The value of land would be lowered by it; the market gardener would be ruined by it. The canal could carry goods cheaper. Steam would vanish before storm and frost; property would be deteriorated near a station. It was called the greatest draught upon human credulity ever heard of. It was erroneous, impracticable, and unjust. It was a great and scandalous attack on private property, upon public grounds — one class was informed that the locomotive would travel so fast that life and limb would be endangered, another was told that it would be too heavy to travel at all.' "[1]

The development of the railway transportation system not only made possible progress in other lines by providing a means for shipping raw materials and finished goods cheaply and expeditiously, but it also stimulated the development of the coal and metallurgical industries, adding as it did to the demand for both coal and iron.

---

[1] Knowles, *Industrial and Commercial Revolutions*, p. 256. Quoted.

Summary of the effects of the Industrial Revolution. The Industrial Revolution gave rise to new industries and brought previously undeveloped industries into prominence. The cotton industry, mining, the iron-working trades, and transportation are examples of this. It caused the break-up of the domestic system of manufacture and its replacement by the factory system with power-driven machinery. New districts in England became important industrial centres. There was a tendency for example for industry to concentrate in the north of England where the iron and coal fields were located. Towns of a somewhat different type, with large industrial populations, sprang up, and from this crowding together of workers in the factories and in the homes developed a spirit of class consciousness greater and more powerful than ever before. The laborer was further removed from his employer than formerly; this too tended to weaken old ties and added to the growing class consciousness. The substitution of machines deprived the laborers of a part of their competitive strength, for it made trade skill and knowledge of less value.

Along other lines great changes were taking place. A new group of capitalist manufacturers was being formed, whose interests lay in an extension of manufacturing and freedom of trade. They wanted as little regulation as possible and cheap food for their manufacturing population. Economic interest and the philosophy of the age worked in the same direction — toward the abolition or lack of enforcement of laws which might hinder the progress of manufacturing. A striking example of the victory of this new class over the old vested interests is seen in the abolition of the tariffs on the importation of grain in the first half of the nineteenth century.

England's productive powers increased by leaps and bounds. In almost every line increased production can be noted, and in some instances the increase was truly phenomenal. A table showing the growth in the output of coal and iron has already been given. In 1801 the export of pig and bar iron amounted to 4,584 tons; by 1850 it had grown to 611,407 tons. In 1780 the official value of cotton goods exported was given as £355,060; by 1800 it had increased to nearly £6,000,000, and by 1850 to over £110,000,000. In consequence of a duty imposed on printed cottons we know the quantity printed until the duty was repealed in 1831.

Year	*Number of yards printed* [1]
1796	20,600,000
1800	32,870,000
1814	124,600,000
1830	347,450,000

Foreign commerce, both exports and imports, showed an almost steady increase. In fact by the middle of the nineteenth century England was in a position of unchallenged supremacy in production and foreign trade.

Paralleling this growth in productive power and in some respects a necessary condition of it, we find a growth in consuming power and a widening of the market. England's manufactures were sent to all corners of the globe and among her best customers was the United States. The United States was increasing in population and in wealth; there were more people to buy goods and they were better able to pay for them. The development in transportation tended to bring this and other markets nearer to the producers, as the cost of shipping goods to and from the ports decreased and as ocean-going steamers began to supply more regular and cheaper service than the sailing vessels.

Producers who had in the past been supplying a fairly steady market were dazzled at the prospect before them and there was a race between producers to see which could turn out the largest volume of goods and reach the market first. They were impelled to this course by the knowledge, which they soon acquired, that they could reduce the unit cost of the good which they were making by increasing the quantity produced.

---

[1] Porter, G. R., *Progress of the Nation*, ed. by F. W. Hirst, 1921, p. 305.

We should not neglect to note that changes in the methods of marketing and of financing manufactured goods were also consequences of the Industrial Revolution. Old methods which had developed to handle a relatively small and stable quantity were found to be inadequate, in the face of continual increases in the output of the factories, and gave way to new methods designed to meet the new situation. A discussion of this phase of the Industrial Revolution must however be postponed until later chapters.

The effect of the Industrial Revolution on the laborers. We should however err were we to leave the impression that the period covered by the Industrial Revolution was one of universal prosperity and contentment. Such was far from being the case, and in many respects the period forms one of the least attractive in economic history.

The hand workers were in a fairly comfortable and contented condition in the first half of the eighteenth century, but by its close conditions had begun to change for the worse. The growth of factory towns made it impossible for the worker to have his plot of ground and do a bit of farming on the side; in fact the workers were so crowded together that the living conditions were miserable. Houses were located in the midst of filth and disease, with practically no attention to sanitation and with little attention to the water supply. The factories were damp, hot, ill-lighted and ill-ventilated, and adequate guards for the machinery were practically unheard of. The competition for employment enabled the employers to exact twelve, fourteen, or fifteen hours of work a day. Women and children were exploited mercilessly; it was not unknown for children of seven to be forced to work at a machine for twelve hours a day, and women were compelled to work as long hours as men at tasks far beyond their strength.

Perhaps a lowering in the standard of living of the workers was inevitable; a period of transition is always difficult, for there is both ignorance of the real conditions and a helplessness in face of the new situation. We can consider the situation which developed as a temporary incident of a great economic change, requiring time for old customs, traditions, and ideals to adjust themselves to new conditions. The prevailing philosophy taught that unrestricted competition was desirable in the public and private interest and doubtless made the readjustment less rapid than would otherwise have been the case. It is well recognized today that public policy demands that the state, for its own safety and health, preserve the health and vitality of the children, but it only gradually dawned on the people of that generation that regulation of the hours and conditions of labor of women and children was necessary.

It should also be noted that the children of the hand workers had led no life of ease before the factory was open to them. As soon as they were able to walk they were set to work by their parents and were expected to earn their living. There was little opportunity for pleasure and none for education in a formal sense. There was almost as much reason to protect the children from the tyranny of their parents as from the exactions and ill treatment of factory foremen. Any factory inspector or employment manager at the present time knows how ready many parents are to lie about the ages of the children, how ready they are to sacrifice health and education for the addition which the children can make to the family income.

The limitation of the number of hours which children could work and the prohibition of factory work before a certain age were forced as soon as the public was fully informed of the conditions, and this created a child unemployment problem, the solution to which was found in state education. In the long run therefore the factory system worked for the benefit of children, but not before it had taken its toll in death and disablement.

For the men too the factory system probably meant an advance, once the period of transition was over. The factory became a better place in which to work than the cottage had been. Factory inspection insured cleanliness, air, and light. The work was monotonous, it is true, but scarcely more so than working at a hand loom, and with a shortening of the hours of labor the discipline of the factory though always distasteful became less burdensome. Work was likely to be more regular than in the domestic system, for the factory owner dared not run the risk of dispersing his labor force during a dull period, and because of his heavy capital investment he could not afford to allow machinery to lie idle if it were possible by any means to keep operating. As time went on workers found it easier to combine into effective unions and force improvement of wages and working conditions. This meant the beginning of labor troubles, but few would deny that the improvement in the life conditions of the mass of workers has been worth while.

The Industrial Revolution has meant not only an increase in the standard of living of the workers but, notably in England, has permitted an enormous increase in population. Formerly England was of necessity largely self-sufficient, but the development of her manufacturing industries in conjunction with the opening up of virgin soil in the New World made possible the exchange of manufactured goods for food supplies and raw materials and thus enabled the population to expand far beyond the limits which would be set by the natural productive powers of the soil. The population of England and Wales, which amounted to about eight and one-half millions in 1801, had doubled by the middle of the century and had increased to almost thirty-eight millions by 1921.

For society as a whole the conditions of life have been improved. The struggle for the bare necessities is not so bitter, and the range of comforts and luxuries attainable by the masses of the people is enormously wider. Less engrossed in the business of making a living than were our medieval predecessors, we are more free for living.

The Industrial Revolution in America. It is questionable whether it is proper to speak of an Industrial Revolution in America. Certainly the improvements in manufacturing which were altering the industrial and social life of England in such a radical fashion had no immediate counterpart in America. The reasons for this are easy to understand. In the first place manufacturing save in the home and for home use was hardly established in the colonial period. The people could more easily secure their manufactured goods by devoting their time to developing their natural resources and exchanging the products of the field and forest for the goods they required. In the second place, the colonial policy of England did not contemplate the development of manufacturing. It was in fact definitely and consistently suppressed. The colonies were conceived to be a source of demand for English manufactures, and not a rival to them. England also tried to maintain a monopoly of her improved machinery by stringent prohibitions on exportation. Hence at the beginning of our national history we find a very great dependence on foreign importation.

The first great stimulus to the development of manufacturing came when the flood of foreign goods was shut off by the Embargo of 1807 and later by the interruption to trade caused by the War of 1812. A cotton spinning mill had been established in Pawtucket as early as 1789 by Slater, who had been an apprentice under a partner of Arkwright and built his machines from memory. But by 1805 only three more had been started. After the embargo other spinning mills were founded, and in 1814 Lowell founded at Waltham, Massachusetts, the first factory using power looms. Following the war there was a period in which the newly established industries were favored by relatively high customs duties, and the manufacture of cotton, woollen, and iron products grew in importance.

The United States did not however offer a favorable field for the rapid spread of the factory system. The richness of her natural resources was such that it was unprofitable to spend more than a small proportion of the available capital and labor in developing manufactures. A practical illustration of this is seen in the difficulty of securing labor for factory work. In England it soon came to be a choice between working in a factory or having no work at all; but in the United States up to fairly recent years there was always the opportunity to secure land in the West at a nominal cost. As a factory laborer a man could expect little more than a modest living; as an independent farmer he might count on a fair living and could reasonably hope for financial independence for himself or his children. Similarly the development of the mineral resources and of transportation diverted capital and labor from manufacturing.

In New England and other regions not endowed by nature with especial richness, manufacturing increased in importance with the growth of population and of capital. Yet if we look for the period in which manufacturing ranks with agriculture in importance we shall find it in the second half of the nineteenth century and not in the first half. We see the results of the Industrial Revolution in America, and America has utilized and improved the inventions of the period, but the growth of manufacturing here has been a gradual process, unaccompanied by the industrial upheaval manifested in England.

Freedom of people
″
Population
″ rich natural resources
Accumulation of capital
[illegible] [illegible] developed faster
[illegible] [illegible] [illegible] [illegible] time
[illegible] of machines

[illegible] review [illegible] also, or more gov't regulation

CHAPTER V

DIVISION OF LABOR AND THE CONTROL OF PRODUCTION

Coöperation. The key to an understanding of the modern industrial system is to be found in the idea of *coöperation.* By this is meant that the divers elements in society are consciously or unconsciously working together toward a common end.

There are historical examples of communities working together and consciously regulating their various relationships to each other. Such societies as the Oneida Community, the Amana Society, Brook Farm, and others were organized for this purpose. Very definite problems have to be settled by the group as a whole or by representatives, as for example what shall be produced, how much of it shall be produced, who shall work at the different tasks, whether one shall be a farmer or a blacksmith, and how the individual members of the group shall share in the benefits of the efforts of all. The decisions finally reached must be accepted by all members of the community, to which end enforcement by authority is usually essential.

Such examples of small isolated socialistic communities are however of only incidental interest to our present inquiry. Our task is to obtain a bird's-eye view of the present-day economic organization in order that we may see in what sense it is a coöperative system and may discover the motive forces which impel the various members of society thus to work together and the regulating forces which direct their efforts.

First of all we must acquire a clear notion of the meaning of coöperation and its several forms. Probably the earliest and certainly the simplest type of coöperation is found where men get together to perform a task which is too great for the strength of a single man. In former years the farmers of New England worked together at a barn-raising or in lifting a millstone into place. There are and always will be in everyday life countless examples of this form of coöperation, which we may call *simple coöperation*, and which is distinguished by the fact that the several coöperating workers are all performing the same kind of work.

Coöperation becomes of vastly greater importance when it passes beyond this simple stage and leads to specialization, in which each worker confines himself to one kind of work, made possible for him because other kinds of work are cared for by other workers. Such coöperation is called *division of labor*. It is of varying complexity according to the degree of specialization. The first steps occur when industry is separated into distinct trades. One man devotes his entire time to making shoes, another to raising wheat, another to making cloth, and so on through the various trades with which we are familiar. But the mere separation of trades, in which each worker performs all the several operations of his own trade, taking the raw materials and finally turning out the finished product, has only begun to explore the possibilities of division of labor. Every trade is itself a complex of many distinct operations, and division of labor has not reached its final goal until each of these operations is made a separate task and assigned to a special group of workers. Although it did not originate there, the more complex division of labor is seen to best advantage in the modern factory, where it has been developed to its highest point. Here we see the extent to which it has been found possible to split a trade up into distinct minor operations, each performed by a different group of workers. In the making of a pair of shoes the soles are cut by one man, the various parts of the uppers by others, one part is stitched by still another, and so on. Complex division of labor may be carried on without power machinery, but its possibilities are usually greatly enhanced by the employment of machinery for many of the operations. In the making of the Goodyear welt shoes it is said that there are one hundred and eighty-five distinct operations, of which one hundred and fifty-seven are performed by machines.

The economic effects of division of labor. Simple coöperation is in fact so simple that it will require no further discussion. Our attention for the remainder of this chapter will be devoted primarily to division of labor as the really fundamental form of coöperation.

The one predominant fact about division of labor is that it increases man's efficiency in production. In the first place division of labor, by permitting each worker to specialize upon one trade or even upon a single operation in a trade, permits him to acquire a degree of skill and dexterity which would otherwise be out of the question. Constant repetition gives a manual dexterity which soon becomes automatic and almost unconscious. Just so far as it is necessary to think out the next move is speed hampered. Contrast the swiftness and deftness of the seamstress, whose needle moves in and out with automatic precision, with the clumsy efforts of the amateur. Contrast the man who is at the same time a farmer, a carpenter, and a blacksmith, with the specialists in these trades. The carpenter instinctively reaches for the proper tool; the farmer has to make a decision, and when he has determined to use a saw, let us say, his hand lacks the cunning which constant use has given to that of the carpenter. The farmer who builds a shed or a barn or paints his house in his spare hours, requires much more time than the professional carpenter or painter. The simplification of the individual operations in a factory makes it possible for the operator to acquire a degree of skill and dexterity in his given operation which often appears little short of miraculous.

Division of labor also makes possible the saving of time which by the unspecialized worker is wasted in putting away the tools required for one job and getting the tools for another. A certain amount of time is used in really getting started, in deciding where and how to begin on the new job. Even a skilled house painter spends some time in getting up momentum, but once the pace is set he can keep it up pretty steadily through the working day. When the day is divided among several different tasks much time can be wasted in the motions of work without accomplishing a great deal. This is one of the handicaps against which farm labor struggles, because of the limited opportunity for division of labor on the farm.

When division of labor does not prevail, there is also the social waste which comes of leaving tools idle during a good part of the day. If analyzed it can be seen that this is also a waste of time for the group as a whole. Assume that we have five individuals and five possible trades. If each individual follows all five trades twenty-five sets of tools are necessary; if there is specialization, only five sets will be required, and they can certainly be made in a fraction of the time required to make the twenty-five, thus releasing labor for other employment.

Division of labor not only economizes the use of tools and machines, but it leads to the invention and utilization of complicated labor-saving machinery which would otherwise not be possible or economical. As the processes of any trade are subdivided each operation becomes more and more simple, and sooner or later it becomes possible to turn over the simpler ones to machines; thus setting free labor for other tasks and increasing the product of industry. The reduction in the physical labor of tending machines has made it possible to use women for a variety of tasks which men alone used to perform, and the decreased difficulty of the task itself has considerably shortened the time spent in acquiring the necessary skill to perform it. Contrast the apprenticeship of seven years with the few days or weeks which seem to suffice now to learn how to operate a machine.

Finally with division of labor it becomes possible for a man to find the work for which he is best fitted and in which he can show the greatest dexterity and speed. The ability to locate the source of trouble in a balky automobile engine seems to be somewhat instinctive, in some measure to be born in a man and not to be acquired. It by no means follows that in our society this adaptation will always take place, and in many cases it is quite apparent that it does not take place. There is a tendency to follow customary lines of work. The son of a banker may follow banking as the line of least resistance, even though his natural bent may be in the direction of medicine. There is many an automobile mechanic who would be a much more useful member of society as a farm hand, and doubtless many a farmer who is wasting his talents on the farm. But there is at least the possibility of this adaptation today, whereas formerly, as on the manor, nearly every man had to be a farm laborer whether he would or not. As a special application of this principle, we have the inestimable advantage of giving full time employment to capable business leaders, whose important place in our economic organization we have already noted. If one will think of some of the great industrial leaders — Carnegie, Harriman, and Hugo Stinnes, to mention only three names — and try to imagine them in the manorial period in any rank of life he chooses, he can better appreciate the extent of the opportunity which the present organization of society gives to native talent and the degree to which society profits from the exercise of such talent.

All of these results of division of labor promote production, by economizing labor, increasing its efficiency, and making more effective use of capital. With a given amount of human labor there emerges a richer flow of commodities and services for the satisfaction of human wants. Or we can look at it from another standpoint and say that division of labor makes it possible to turn out the same quantity of goods with less expenditure of effort, thus leaving more leisure for the enjoyment of the products of industry. It should be noted however that not everything that can be said of division of labor is in its favor. It is objected for example that modern specialized labor is monotonous and deadening in its effects on the workers, that modern machine tending is so largely automatic that skill is at a discount and labor is kept at a dead level, that the highly specialized laborer is unfitted for work in another trade. There is doubtless some measure of truth in these charges, but they may easily be exaggerated. Labor in the medieval manor was hardly free from monotony, and the shorter working hours of modern industry probably turn the scale in favor of the present day as regards the deadening effect of labor. For machine labor it can be said that, automatic though it be, the demand for skilled labor is still active and that labor could scarcely be more nearly on a dead level than it was on the manor. The charge that specialized training today makes change of occupation difficult is the reverse of the truth. An operative in a watch factory could fit into a bicycle factory more easily than a house painter could become a blacksmith. On the whole such drawbacks as are inherent in the division of labor are certainly not sufficient to prevent its net effect being overwhelmingly on the credit side of the account.

Division of labor and the extent of the market. Whether the farmer builds his own barn in addition to farming or hires a carpenter to do it for him depends in part on conditions over which he, as an individual, has little or no control. Assuming that his true bent is farming, he will prefer to hire the carpenter, but if he cannot sell the wheat that he does not need for his own family, he will have to build his barn. If the carpenter cannot sell his services he is forced to raise his own wheat, even though he be an indifferent farmer. The possibility of exchanging the surplus which one individual produces is a necessary preliminary to division of labor, and as exchange becomes easier there is a tendency to subdivide the tasks and to proceed from simple to complex division of labor. On the manor division of labor was of the most rudimentary sort, because there was no market in which surplus products could be sold. When the towns began to develop and to provide a market we find the beginning of specialization on the part of the town craftsmen and the country laborers and the exchange of the surplus of each. When the market consists only of the town and the countryside, the village cobbler has little use for assistance and can himself fabricate the whole shoe. If he does require one or two helpers there will be little division of labor among them. Each will be able and will be required to make a whole shoe, although it may develop that one will be particularly skilful in a special part of the work and devote a large part of his time to that operation. But if the shoemaker is called upon to manufacture several thousand pairs of shoes a year, it will be economical for him to hire men to perform minor operations, for then he can provide them with continuous work. It would further be profitable for him to use machinery to perform some of the operations.

Territorial division of labor. The extent to which division of labor will be carried depends then on the extent of the market, on the volume of goods which consumers are willing to purchase. The demand for a particular product may increase with the gradual growth of the population, or as the result of a change in popular taste. But in the past the most striking cause of increased demand for the product of a particular manufacturer or locality has been a widening of the market through improvements in the means of transportation; in other words, through the addition of a demand which was previously satisfied by other producers or not satisfied at all.

Certain regions are peculiarly fitted by natural endowment for the pursuance of certain occupations. The soil and climate of the American Southern states make them particularly suitable for the growing of cotton; the coal and iron of Pennsylvania make specialization in iron and steel products natural; the fertile plains of the mid-western states offer unusual advantages for growing grain. The specialization in production which is thus dictated is called *territorial* or *regional division of labor*. Under almost any circumstances there will be some degree of specialization in the products of the region, but unless adequate transportation facilities exist between different regions each must be self-reliant to a great degree. It is a familiar fact that a century ago New England was in large measure dependent on her own food supply, in spite of the fact that her soil has never been as productive as that of the West. Yet New England could not specialize in manufacturing until she could market a volume of goods far in excess of the needs of her own people and could satisfy her food requirements by grain from the West. The great farming regions of the West, on the other hand, would never have been able to specialize in agriculture except for the possibility of selling their products in New England and other regions and purchasing from New England and other regions their stock of manufactured goods. It was the development of the transportation system that made possible this interchange of goods and this territorial division of labor, to the great advantage of both regions.

Division of labor a form of coöperation. Production with division of labor is clearly coöperative. Laborers, employers, capitalists, and landowners are working together in the sense that each is in part dependent on the efforts of others. In the long chain of productive operations each is dependent on the producer below him for his raw materials. The cotton manufacturer relies on the grower of cotton, the cotton middleman, the coal miner, the manufacturer of cotton machinery, and literally a host of others, to provide him with the wherewithal to conduct his operations. The agriculturist grows more wheat than he can use, confident of his ability to sell the surplus and with the proceeds secure the cotton cloth that he needs. Coöperation in production consists in the creation of a surplus and the exchange of that surplus for the production of others. The laborer working for wages just as truly exchanges his surplus as does the manufacturer who hires him, for the manufacturer pays him the present worth of his product in money instead of giving him a portion of the product to sell, as might be done and as is sometimes done with farm labor.

This fact of coöperation is by no means of academic interest only; it is of great practical importance. In the past material progress has followed in the steps of coöperation. Each one of the industrial stages which we studied earlier is marked by a widening of the coöperative area and by an increase in the output of goods. As in the past so in the future, one path of material progress lies in the direction of increased coöperation, of greater specialization both within the industry and within countries having natural aptitudes in the production of certain goods. Any artificial restraint therefore which hinders this specialization decreases by so much the effectiveness of labor.

We cannot leave this topic without emphasizing the reliance which individuals, communities, and nations place on the effectiveness of coöperation in supplying them with the goods in which they are deficient. England affords the classic example. Lacking in a sufficient food supply and in raw materials such as cotton, copper, and many minerals, she has been able, through utilizing her deposits of coal and iron and her intelligent population, to produce a volume of manufactures great enough to supply her own needs and at the same time, by selling the surplus abroad, to procure food for her population and raw materials for her industries. Any large city of the present day is entirely dependent on outside sources of supply for food and raw materials and is therefore dependent for its very life on transportation facilities. The tying up of all the rail connections to New York would be a catastrophe of the greatest magnitude. It can easily be seen that the threat of a nation-wide strike on the part of the railway workers is a powerful weapon, and that it is powerful because of the degree to which we depend on coöperation in production.

The control of production. When we survey production as a whole we see the truly marvellous spectacle of millions of men engaged in thousands of different tasks, turning out quantities of goods varying in kind from the most essential to the most frivolous, transporting these goods where they are most urgently wanted, financing their manufacture, shipment, and sale, and performing personal services of the most varied character. Each producer is apparently quite independent in his choice, yet each seems as a rule to turn out exactly the amount of his particular product which can be sold, and with almost no exceptions the goods all seem to reach the consumer where and when they are wanted. This fine coördination of human activity cannot be the consequence of accident; such marvellous coöperation is not merely a fortunate coincidence. There must be some control. Now control may be of two types: (1) by ~~legal authority restricting the freedom~~ of the individual producer, or (2) by ~~economic forces~~ operating under a régime of personal liberty. In a certain sense we may call the first conscious, the second unconscious, regulation. Of the two types of control it is evidently the second which governs in the modern world. But this has not always been so.

Regulation by legal authority. For the sake of a background for our study of the modern regulatory system let us digress for a moment and revert to a time in England when it was widely held that regulation of individual activity should be a conscious and definite policy of the government. The motives for regulation were in part to protect the subjects against the dishonesty or oppressive power of other nationals or foreigners and in part to increase the prosperity and prestige of the nation.

Protection for the individual. We have already seen that in the period of the town economy the protection of the inhabitants of one town against those of another was, in the modern view, carried to extreme lengths. The fixing of the price of bread was a common practice and in the same spirit laws were passed providing for the inspection of cloth and other articles. This meant the practical determination by authority of the type of cloth which could be made. Some of these regulations lasted well into the nineteenth century, when they were either repealed or ceased to be enforced.

Another form of regulation was interference with the freedom of movement of the individual and limitation of his choice of occupation. We have already referred to the way in which the manorial laborer was tied to the soil. In the course of time the absolute restraint on moving from place to place was relaxed, but it was still true in the seventeenth and eighteenth centuries that custom and law made the laborer, and particularly the poor laborer, relatively fixed to the place in which he was born.

In the reign of Elizabeth the law of apprenticeship was passed, which compelled any person who wished to follow a trade to serve an apprenticeship of seven years. This simply followed the custom of the time and was intended to make the efforts of the gilds to ensure honest workmanship more fruitful, but it tended to solidify a custom which might more easily have been modified in the course of time but for the law. It is true that this law applied only to such trades as were in existence when the law was passed and applied only to towns, but municipal regulations made it difficult for anyone to follow a trade unless he had been an apprentice for the required length of time.

Wages also were regulated by authority. According to the act of Elizabeth referred to above and subsequent legislation, the justices of the peace in each jurisdiction were required annually to determine what wages should be paid. It may be said in this connection that their efforts were not entirely unwelcome to the laborers, for while the justices were undoubtedly inspired with conservative views, it is nevertheless true that they frequently raised wages. In a period when workers were scattered and of negligible importance politically, concerted action on their part to raise wages was quite out of the question. Hence any body of men which might even slightly consider their needs was not without advantage to them.

Promoting the national welfare. The second motive leading to regulation was the desire to increase the prosperity and power of the state. The sixteenth and seventeenth centuries were the ones in which the nationalistic spirit was manifesting itself to an ever increasing degree, and it was generally agreed that the first duty of a ruler was to expand his power whether at the expense of foreign nations or of his own subjects. Inasmuch as it was further believed that national and private interests were usually incompatible, it is easy to understand the restrictions placed on the private individual.

In furtherance of the aim to promote the national prosperity and power much attention was given to the regulation of the nation's foreign trade. Armed conflict was an almost constant feature, and the employment of paid standing armies in place of the feudal levies of former times as well as the increase in the scope of the state's activities made the accumulation of a store of the precious metals highly desirable, for in time of war that state which was most plentifully provided with ready money enjoyed a great advantage. This led naturally and almost inevitably to a belief that gold and silver were the most important forms of wealth, and measures were taken to conserve and increase the supply. The most obvious and the earliest of these measures was a prohibition of the export of the precious metals. This legislation was unsuccessful in operation and was soon seen to be unwise, as there were occasions on which the export of gold and silver might result in an augmented importation.

Public attention was then directed toward starting a flow of gold and silver toward England through the regulation of imports and exports of goods. It was held that if the total value of the exports was greater than that of the imports the balance would of necessity be paid in money. That condition was called a "favorable balance of trade," and an "unfavorable balance of trade" was said to exist when the value of the imports exceeded that of the exports and money flowed out.[1]

[1] It is not our purpose here to point out the merits of this theory, but there are so many misconceptions based on it even at the present day and it is so widely held in one form or another, that it must be stated emphatically that neither an excess of

In order to secure a favorable balance of trade and to avoid an unfavorable balance heavy duties were imposed on the import of manufactured goods or on goods which could be produced in England. In some cases bounties were granted on the export of English wares. The importation of raw materials which could be worked up in England was generally encouraged, for the value of the goods when exported would be greatly enhanced and would therefore help swell the value of the exports. So also the export of domestic raw materials which could be manufactured at home and bring a much greater price in the finished state was prohibited. This was true of wool in particular. In summary we may say that regulation was intended so to affect the course of foreign trade that England would import only raw materials which could later be exported in a more valuable form and would export only finished goods.

Mercantilism. The political and economic philosophy of the period, which embraces these regulatory provisions, is called Mercantilism. The foregoing account makes no pretense of giving a complete picture of all the restrictions thus imposed upon the producers by the laws of England, which, by the way, were distinctly more liberal than those of most other countries. As compared with the conditions which prevailed in France for example English workers and traders enjoyed a great measure of freedom. England had free internal trade by the eighteenth century, whereas the internal trade of France was impeded by innumerable customs barriers, and in the period of most minute regulation the individual could not determine what he should produce, how much he might produce, where he might produce it, or to whom he might sell it.

Laissez faire. During the eighteenth century the conviction spread that there had been too much government, too much interference with private business, and there arose a strong reaction against the restrictions of mercantilism. *Laissez faire* — let things

exports nor an excess of imports can be considered as either favorable or unfavorable and that it is impossible over a period of time for a country to have either a continuous inflow or a continuous outflow of gold, unless one of the countries is a gold-producing country and is exporting the gold in the same fashion as the United States exports wheat. This subject will come up for full discussion in a later chapter.

alone — became the watch cry of philosophers, economists, and statesmen. *Laissez faire* in the first place emphasized the importance of the individual and of individual welfare. Secondly it assumed that individual welfare and national prosperity are not incompatible. It went further and asserted that individual prosperity is at the very foundation of national prosperity; that the individual, if allowed to follow his own self-interest without external restriction, will exert himself to the utmost in the line of endeavor for which he is best fitted and will therefore be a more productive and hence a more valuable member of society than if his occupation and his conduct in that occupation be prescribed for him by either law or custom. As each man follows his own interest the clash of conflicting interests will lead to modifications and compromises which will result in approximate justice to all.

Translated into terms of concrete policy, *laissez faire* supported (1) free trade, as opposed to protective tariff, (2) no attempt to fix or regulate prices or to regulate either the quantity or quality of goods produced, (3) entire freedom on the part of the individual to enter any occupation he might choose, in any place he might think advisable, (4) the determination of wages by unfettered bargaining between employer and employee, (5) the limitation of the functions of government to (a) defense and the maintenance of peace and order, (b) the dispensing of justice, (c) the establishment and maintenance of certain public works and institutions, such as the paving of streets, the improvement of harbors and means of communication, the care of education, etc., which cannot be made commercially profitable as private enterprises.

The progress of liberal ideas. The reaction against mercantilism spread rapidly in the eighteenth and nineteenth centuries, and the general acceptance, particularly in England, of the doctrines of *laissez faire* ushered in a period in which the dominant note was the absence of conscious regulation and the free play of the forces of individual initiative. The basic principles of *laissez faire* have on the whole successfully maintained themselves to the present day, though experience has shown untenable some of its extreme positions, both of philosophy and of practical policy.

The present system of economic freedom. The conditions under which production operates at the present day are on the whole those of economic freedom in contrast with legal authority. Taking a broad view of the present system, we note that it is founded on three great principles: (1) personal liberty, (2) private property, and (3) individual initiative and control of enterprise.

Personal liberty. So generally is personal liberty today taken for granted that we are in some danger of failure to appreciate its full significance. In truth personal liberty is a comparatively new thing, only recently gained as the climax of a struggle of centuries. Throughout most of the world's history freedom was a rare privilege enjoyed by the favored few. The masses of the people were slaves or serfs, more or less completely subject to the will of their owners or overlords. Even after slavery was abolished and the status of serfdom gradually liberalized and finally developed into comparative personal freedom, the legal restrictions upon economic activity which we have been studying continued. Real economic freedom is scarcely a century old.

Private property. In Chapter I, where property was defined and analyzed, the institution of private property was taken for granted. But this also is a remarkable and a fairly modern institution. There was probably very little private property in ancient days, common ownership by the tribe or nation or ownership of all things by the chief or king being the rule. The feudal system was based upon ownership of at least the most important element of the wealth, the land, by the king or other ruler. Even today a certain part of the wealth of every community, more in some countries than in others, is owned in common as the property of government. But in most modern nations the great bulk of all the wealth is owned by individuals. Private property involves not only the right of the owner to enjoy all the benefits of his wealth but also the right to say who shall own it after his death; that is, the right of bequest or inheritance. This attribute of private property greatly increases its significance as one of the conditions which control production in the modern world.

Individual initiative and control of enterprise. The development of this institution, succeeding the régime of control by legal authority, has already been sufficiently discussed. Today it is not the congress or the legislature, the town council or the gild authorities that say what sorts of goods shall be produced and what quantities of each, where they shall be marketed, and at what prices sold. All such matters are left to the personal judgment and free initiative of the men who direct modern industry, the farmers and miners, manufacturers and merchants, bankers and brokers; in short, the entrepreneurs.

The motive to productive effort. The goal of all industry is production, in order that human wants may be satisfied. Each producer is, under the institution of personal and economic freedom, at liberty to produce whatever he chooses, and the institution of private property permits him to have and dispose of what he produces. Thus is the desire for the satisfaction of wants liberated to become the motive force behind production, as it could never be without freedom and private property. The slave may cherish wants without result in increased application and diligence. He does as little as he can without incurring the wrath of the overseer, because he knows that there is small connection between his efforts and the satisfaction of his wants. A freeman, on the contrary, has the right to enjoy the fruits of his toil, and he is spurred on to efforts more or less energetic and effective according to the urgency of his wants and his productive capacity.

But, as we have seen, modern society is organized upon a coöperative basis, which means in the first place that one does not produce the things that will satisfy his own wants but must first exchange his own product for goods produced by others and secondly that production must submit to an extremely complex organization controlled by certain individuals, the entrepreneurs. For the great mass of the workers, the connection between effort and satisfaction of wants, though it still exists, becomes less direct and less obvious. Nevertheless the basic motive remains; the wage earner is aware that the amount of his wages depends, in some measure at least, upon the efficiency of his productive efforts, and the entrepreneur thinks of the profits of his business. The modern motive to productive effort is thus frankly self-regarding, and modern industry proceeds along the lines of competition, which means in general that each is acting independently in his own interest without regard to the interests of others.

What guides the entrepreneurs? Such being the motive to production, and its direction being in the hands of a comparatively few entrepreneurs, what is it that guides the judgment of the latter and tells them what goods to produce, what quantities of each, where to send them, and to whom to sell them?

The answer is price. Prices determine profits and invite or repel the entrepreneur as he casts his eye over all the possibilities in the field of production. Wages, the price of labor, turn laborers from one industry to another. Price directs the flow of savings into new capital equipment for this industry or that. Price determines the amount of iron ore that will be mined, the amount of wheat that will be raised, whether wheat will be sent to one country or another. Price is so omnipresent and so subtle a force that few are fully aware of its significance or realize how completely it controls our economic life.

A simple illustration will make this clear. Ask any farmer what he would do if he knew or firmly believed that the price of wheat were going to double. Having every reason to anticipate greater profits, he would undoubtedly try to grow more wheat per acre and plant some acreage to wheat that had previously been devoted to another crop or lain fallow. In his efforts to produce more wheat he might bid for laborers, offering them higher wages to induce them to leave their present jobs, and he might purchase new machinery, so stimulating the manufacture of farm machinery and directing capital into that line of industry. Conversely a fall in the price of wheat, threatening lower profits or even losses, will reduce its production, may lower farm wages, and tends to check the flow of capital into farm machinery. The farmers of this country rejoiced in the high prices of the World War. Profits were large and production expanded enormously to make up for the interruption to the flow of wheat from Russia, Australia, and the combatant countries. The return of peace brought with it an approach to normal conditions. Australian wheat could get to the market, and France could devote capital and labor to raising wheat. The price of wheat dropped and American farmers found themselves in a desperate plight. The explanation was simple; the world production of wheat was too great to be sold at a profit by all of the farmers. The automatic check was working, painfully no doubt, in the direction of decreased production, the only way to raise the price to a remunerative level.

Price is thus the medium through which the consumer makes his voice heard, telling those whose judgment directs industry how much of each good to produce and whether any particular good shall be produced at all. If someone is willing to pay a sufficient price, anything conceivable will be produced, from bread and shoes to diamond shoe buckles. If for something else no one is willing to pay a price from which the producer can make a profit, then that thing will not be made. There is no production of the old-fashioned spinning wheels or of clothing of a bygone style or of gold horse shoes, because no one will pay the price. High prices of the products in any particular line attract the attention of the entrepreneurs and they direct land, capital, and labor into that branch of production. Low prices are the automatic check to overproduction. Prices of raw materials and wages govern costs and have the opposite effects, stimulating production when low and putting on the brakes when high.

What is back of price? Price both controls production and is itself controlled by production. The full explanation of the economic mechanism which determines prices must be postponed to later chapters, but the reader can readily understand that a great stock of any particular product inclines the producers to offer to sell at lower prices, while a shortage causes prices to rise. We thus complete the circle of cause and effect. A surplus of any particular good relative to goods in general, by making its price low, tends to check further production till equilibrium is restored; a shortage of any given product tends to a rise in its price and so stimulates its increased production. The reader will also recognize that this mechanism by which price regulates production and is itself regulated by production is an automatic thing not consciously imposed by any person.

It is possible to conceive of a system whereby prices are fixed by legal authority. Let us suppose that the government undertakes to regulate the price of wheat. In some way it must be provided that the right amount of wheat be produced and that the use of wheat be so regulated that a given harvest will last until the next crop is harvested. Now it is possible for the government to buy up all the wheat and dole it out, so much per person, at a fixed price, and thereby see that the stock does last. What is more likely is that the government will simply fix, from time to time, the price at which wheat may be sold. If the price thus set is low, consumption goes on at a rapid pace, the total stocks of wheat may be prematurely consumed, and production of the future may be seriously curtailed. If the price is very high, higher than would be fixed without government interference, the stocks of wheat are kept unnecessarily large, and an oversupply is likely to be produced for the future. Political complications are also likely, either from protests of the farmers if the price is thought too low or from the complaints of the general body of consumers if they consider the price too high.

Returning from this digression, we note that price is actually determined by bargaining between sellers and buyers, producers and consumers. Each consumer of wheat is striving to purchase his wheat at the lowest possible price, while each seller strives to sell for the highest possible price. The only restraint imposed is that of self-interest. A seller hesitates to be too grasping lest he lose his trade to another seller; a buyer fears to hold off for a low price too long lest he lose the opportunity to buy at any price. If the existing stocks appear deficient, this bargaining of sellers and buyers tends to a rise in price and economy in present use with more held back for future needs. The opposite results follow the fall in price which is induced by abundant present stocks. And finally the high or low price encourages or discourages future production to replace the present stocks.

Unconscious coöperation. The modern system of production is coöperative, but it is unconscious coöperation, in that the agents of production are following their own separate interests — the securing of the highest possible wages, or the largest profits possible, or the maximum rate of interest, all forms of price — and are rarely aware of the fact that they are engaged in coöperative enterprise. Indeed under the complex conditions of modern life conscious coöperation through the control of legal authority is scarcely conceivable. To coördinate the supply of goods with the demand for goods, to determine the basis of exchange for all goods and services, to fit workers into their proper niches, to attend to the thousand details which individual interest now takes care of, is a task far beyond the organizing ability of the human race in its present stage of development.

CHAPTER VI

THE FORM OF THE BUSINESS UNIT

As was disclosed in the chapter just preceding, the control of production in the modern world has been entrusted to the entrepreneurs, who are permitted to organize and direct industry, impelled primarily by the motive of private gain in the form of business profits. In the course of the extraordinary era of economic progress ushered in by the Industrial Revolution, old forms of the business unit have been modified and new ones introduced in order that the several branches of industry might adapt themselves to the varied and changing conditions and might function more easily and more efficiently. The form of the business unit is a matter of very considerable significance.

The individual proprietorship. A business which is simple in nature, which requires little capital, and to which attaches relatively little risk does not require a complex form of organization. In fact complexity of organization would prove more a handicap than a help. The small retail grocer or baker for example generally finds the individual proprietorship best suited to his needs. In this form of organization the proprietor is in sole charge of the business, responsible alike for its success or its failure. He may start up in business when he chooses and cease operations when he so desires, without the bother and expense of legal formalities. Unless an activity is specifically prohibited by law, such as the coining of money or the carrying of mail, no line of business is closed to him. Responsible to himself alone, he is in a position to make decisions at once without seeking either the advice or assent of an associate, and he is thus able to take advantage of the opportunity of the moment. His business secrets cannot become common property through the indiscretion of associates.

While there are advantages for the small business in this form of organization, there are certain drawbacks which make it undesirable for a concern of any great size. In the first place the individual is seldom able to command a large amount of capital, for he is rarely in a position to invest as much in his own business as can be secured by a partnership or a corporation. And if he is able to do so, the great risk involved is a deterrent, for he is personally liable for all debts of his business. Hence at the present time we find relatively few large enterprises organized and operated by a single individual. In agriculture alone the individual is still the predominating figure, and, while the number of manufacturing plants owned and operated by individuals is over one half of the total number, the output of these plants forms an insignificant part of the total output.

The partnership. The partnership is a more complex type of organization than the individual proprietorship, for it involves relations between the individual partners as well as the relationship of the partnership to the public. Essentially it is a voluntary association of two or more individuals for the performance of a specific or a general object. This object may be accomplished in a day or a week, or the association may be formed for more general purposes, such as the establishment and operation of a factory, and may last a longer time.

In any case the purposes for which the partnership is formed, its duration, the ways in which it may be dissolved, and the mutual rights and obligations of each partner to the other will be the subject of an agreement usually set forth in written form. As was true in the case of the individual proprietor, the partnership can be created with little legal formality and without delay.

While the partnership agreement controls the relations between the individual partners and can determine the share of profits going to each partner and the proportion of the common losses which each must pay, it cannot control the relations with the general public. From a legal standpoint each partner is considered to be an agent of the partnership, fully empowered to bind the partnership in all lawful undertakings. In the event of inability to meet all debts from the funds of the partnership, each partner is liable to the full extent of his personal resources for the debts of the business, even though the agreement may make one partner responsible for only a certain proportion. For example, let us assume that the partnership agreement of Jones and Brown stipulates that Jones shall receive seventy-five per cent of the profits and be responsible for seventy-five per cent of the losses. Large losses occur and the debts exceed the resources of the partnership. Jones finds himself unable to meet his share of the losses, and therefore Brown must make up the difference if he is able to do so. This is known as *unlimited liability* and is one of the chief points of difference between the partnership and the corporation.

The fact that there are several men united together makes it possible for the partnership to control a larger amount of capital than can the individual. In fact many partnerships have been formed when the individual had reached the point in the development of his business where additional capital was necessary and could most easily be secured through the admittance of an associate. Few concerns are able to finance themselves with their own capital exclusively; they must rely to some extent on bank credit, secured through loans from banks, and on commercial credit, obtained by purchasing goods to be paid for in the future. The ability of the partnership to secure capital in these ways is likely to be greater than that of the individual, in so far as the unlimited liability of several men offers a greater margin of safety to the creditor than the unlimited liability of one. Just so the credit standing of a partnership having about the same net assets as a corporation is likely to be better than that of the corporation, because of the unlimited liability of the partners as contrasted with the limited liability of the stockholders of the corporation.

Another advantage of the partnership over the individual proprietorship may come from the association of abilities. An important question will be viewed from more than one point of view, and the decision which is reached may therefore be sounder. Offsetting this there is the possibility of serious disagreement and consequent delay in making decisions, but at least snap judgments are likely to be avoided. A certain degree of specialization is possible, as in a wholesale house where one partner is entrusted with the buying, another with the selling, and the third with financial questions. This feature is especially prominent in partnerships of professional men, such as those formed by several doctors, each of whom may be a specialist in a particular branch of medical practice.

One of the chief weaknesses of the partnership arises because of its enforced dissolution on the death of a partner. Under the common law doctrine,[1] unless modified by statute law, the death of one of the partners automatically dissolves the partnership and makes necessary an accounting and distribution to his heirs of the deceased partner's share of the net resources of the partnership. This may be a hardship, both to the partnership and to the heirs. But the heirs cannot join the partnership without the formation of a new partnership.

The theory on which this doctrine is based is that the partnership is a peculiarly personal relationship and that anything which occurs to interfere with this must necessarily bring the original arrangement to an end. A logical corollary is that a partner may not transfer his interest in the partnership without causing a dissolution of the partnership, because no associate can be forced on the other partners against their will.

But the partnership has such obvious advantages over the individual proprietorship that, before the corporate form was devised and made common, it was the prevailing form of organization in enterprises of moderate and large size, and it is still not uncommon to find large concerns conducted for many years under a partnership agreement, although the corporation has steadily increased in favor and now overshadows the partnership in importance.

The limited partnership. A special type of the partnership is known as the *limited partnership.* This involves one or more general partners, who manage and direct the business and have unlimited personal liability, and one or more limited partners, whose liability is limited to the extent of their investment in the business. While the right to associate in business as partners without special sanction is well recognized in our law, the formation of a limited partnership is a privilege conferred by statute law, and certain formalities are prescribed, such as registration and the giving of public notice of the limited liability of certain partners. It should be noted furthermore that limited liability does not imply freedom from liability; it means rather that in case of losses the personal resources of the limited partner are not all available for the settlement of the debts of the partnership. Usually the liability is limited to the investment in the business.

---

[1] Common law is based on universal custom which has received recognition and sanction in decisions rendered by the courts. Statute law consists of laws enacted by a legislative body, such as the Congress of the United States or the legislature of one of the states, and enrolled on the statute books.

Another characteristic of the limited partnership is that those partners with limited liability must take no share in the control of the business. If their relation to the business is such as to lead the public to believe that they are actively engaged in the business, the plea of limitation of liability may not be accepted by the court in settling the affairs of the partnership.

This form of organization may prove more advantageous than the general partnership or even the corporation. It offers a method whereby capital may be secured without relinquishing control of the business. A partner, anxious to withdraw from active business, may do so, leaving his share in the business intact. A partnership which has been broken up by the death of a partner may be re-formed as a limited partnership, without disturbing the assets or breaking the continuity of operations, by admitting the heirs as limited partners. In states where the expenses of incorporation are heavy, where state supervision of corporations is quite strict, or where corporations are heavily taxed, the limited partnership may be a more desirable form.

The corporation. Though the most significant type of business unit in the United States at the present time, the corporation is of comparatively recent development, the great increase in the number of corporations having come since the Civil War. The importance of the corporation can easily be seen by the fact that, although only thirty-one per cent of all the manufacturing establishments in 1919 were owned by corporations, they turned out almost eighty-eight per cent of the total value of the products, and in certain industries we find the corporation controlling an even greater percentage of the output. For example, in the manufacture of rubber boots and shoes corporations have entire control. They turned out ninety-seven per cent of all the agricultural implements made in the United States and ninety-nine per cent of the clocks and watches. In practically every line where the market for the product is more than a local one the corporation is found to hold a preëminent place.

The nature of the corporation. A corporation has been defined in part as "an artificial being, invisible, intangible, and existing only in contemplation of law." This definition indicates one of the most distinctive characteristics of the corporation; in the eyes of the law it is a legal or artificial being, a fictitious person, conceived to have an existence apart from that of its owners. The partnership exists in the persons of the partners. For example, suits by or against the partnership are brought in the names of the partners, but the corporation sues and is sued in the corporate name.

Furthermore the corporation depends for its existence upon the sanction of the law. The individual business man and the partnership carry on business by virtue of rights long recognized by common law, but the corporation exists only through the official act of the state. In the United States a special act of the state legislature was formerly necessary for the formation of each corporation, but there are at present general corporation laws which prescribe the method of incorporation.[1] If the organizers fulfill the requirements as laid down in the corporation laws and make application in due form to the proper state official, a *charter*, or *certificate of incorporation*, will be issued to them defining the rights and duties of the corporation.

Called into being by the act of the state, the corporation exercises only such powers as have been conferred upon it. A partnership may engage in any activity which is not prohibited, but the corporation is restricted to the range of activities sanctioned by the terms of its charter or the general laws, either explicitly or implicitly, as essential to its existence. Acts beyond the scope of the corporation's powers are termed *ultra vires* and may result in the revocation of the charter and the dissolution of the corporation.

---

[1] Some corporations, such as the national banks, are chartered under federal laws.

The doctrine of personal relationship, which is stressed in partnerships, is not so important in the case of corporations; the association of capital is of greater significance than the association of individuals, and this finds its exemplification in the right of each of the owners of the corporation to transfer his property rights in the corporation at will, without seeking the assent of the other owners, either individually or collectively. It follows that the death or disability of one of the owners of the corporation does not bring about its dissolution. This makes the corporation a more permanent form of organization.

In some cases indeed a corporation has perpetual existence, but it is customary at the present time to limit the life of a corporation to a definite period, such as twenty or fifty years. At the expiration of this period the corporation is automatically dissolved, but renewal of the charter is so easy to obtain that there is virtually perpetual existence. Dissolution may also result, as we have indicated above, by *ultra vires* acts of the corporation. As a matter of fact however the authority granted by the charter is usually couched in such general and comprehensive terms that the ordinary corporation is in little danger of exceeding its powers. Finally dissolution may be brought about by the action of the majority of the owners with the assent of the state authorities.

Another distinctive feature of the corporation, as compared with the partnership, is the *limited liability* of the owners of the corporation for corporate debts. This is not a necessary mark of the corporation; there have been corporations the owners of which have had full personal liability for corporate debts, but it is usual at the present time to limit the liability of the owner to his investment in the business. A man who has purchased 500 shares of stock in a corporation is liable to lose what he paid for them, but the law does not lay hold of his personal resources to satisfy the creditors of the corporation. In some cases there is *double liability.* The purchaser of $100 of stock in a national bank for example may not only lose the $100 which he has invested but may be required to contribute $100 more to meet the debts of the corporation.

In concluding this brief survey of the characteristics of the corporation we should note that the management of the affairs of the corporation is delegated by the owners to hired managers. The owners elect representatives from among themselves, called *directors,* to care for their interests and to supervise in a general way the operation of the business, but the active management of the corporation falls upon the officers, who are elected usually by the directors. Problems of importance are referred by the officers to the directors; questions of policy or of unusual importance only are submitted to the owners collectively for action.

Capital stock. The capital stock represents ordinarily the original and permanent investment of the owners in the corporation. Suppose that a group of men incorporate the Steel Manufacturing Corporation and agree to contribute $1,000,000, divided into 10,000 shares of $100 each. A takes, let us say, 5,000 shares and pays $500,000 into the treasury for them; B takes 1,000 shares, and so on until all the shares are sold. The corporation now owns $1,000,000 in cash, and the organizers of the corporation have certain property rights against the corporation. The collective name for these property rights is *capital stock,* and if we say that the capital stock of the company is a million dollars, it is an elliptical way of saying that the property rights which the owners of the corporation hold against the corporation amount to a million dollars.

Each purchaser of a share of stock, or *shareholder,* becomes a part owner in the corporation, and as such is possessed of certain rights against the corporation, such as the right to receive dividends when legally declared, the right to share in the resources of the company if it is liquidated, and usually the right to vote at the meetings of the shareholders. A *certificate of stock,* signed by the proper officials of the corporation, serves both as evidence of the shareholder's rights of ownership and, when properly endorsed, as a convenient [illegible] greater capital

method of transferring his rights to another. It must be emphasized that a stock certificate is not synonymous with a share of stock; a share is one of the units into which the total property rights of the owners have been divided, and the certificate is merely presumptive evidence that a part of the total rights is vested in the holder. This is an example of the distinction between a property right and the document in evidence thereof which was pointed out in Chapter I.

Many corporations find that their financial needs are best met by the issue of two types of capital stock, which differ from one another in some important respects. These are *common stock* and *preferred stock.*

Common stock usually carries with it the exclusive right of voting at shareholders' meetings. Preferred stock also may confer the right of voting, but usually the preferred stockholder has no voice in the management so long as his dividends are regularly paid. The ownership therefore of more than one half of the outstanding common stock will give control of the corporation, since the owner has one vote for each share of stock he holds. In fact in a large corporation a much smaller percentage will usually suffice because of the lack of interest on the part of a number of the small stockholders, and of division among the larger ones.

Dividends are declared on common stock by action of the board of directors when profits have been earned and when it is judged wise to distribute them, and the amount or the rate of the dividend is entirely within the discretion of the board. Hence the stockholders in a new corporation are uncertain of a return on their investment for some time, conservative policy dictating that the earnings be reinvested in the business, but as the corporation becomes more and more successful the dividend rate rises until it may be much higher than on any other type of security.

Preferred stock differs from common stock in that a fixed rate of return is promised, and in view of this the rate of return is somewhat lower than that which can be expected eventually from common stock. Furthermore dividends must be paid on preferred stock before any can be declared on common stock. Ordinarily the preferred stockholder has no redress if the company is unable to pay dividends in any given year, and if during the succeeding year very large profits are made, he does not then receive dividends for the previous year. But there is frequently *cumulative* preferred stock, upon which back dividends must be paid up before any dividends can be declared on the common stock. This protects the preferred stockholder against manipulation of the accounts to conceal profits and avoid thereby the paying of dividends. Preferred stock may also be *participating*. This means that when a dividend has been paid to the common stockholders at the same rate as to the preferred stockholders, both the common and the preferred stockholders share in proportions already determined in any further dividend payments made for that dividend period.

The owner of this type of stock has preference over the common stockholder in another direction also, *i.e.*, in the distribution of assets in case of liquidation. The common stockholder receives his share of the assets only after the preferred stockholder has been paid in full. If the remaining assets are insufficient to redeem the common stock in full, the amount is divided pro rata among the shareholders; if there is more than enough to redeem the common stock, then the remainder is usually divided pro rata among the common stockholders.

Bonds and notes. A part of the funds for the operation of the corporation are often secured through the sale of its bonds and notes. A bond is a certificate of indebtedness issued by a corporation, bearing interest at a stipulated rate and payable at stated times and having a certain number of years to run before the corporation is obliged to redeem it. Essentially it differs from a promissory note only in the length of time before maturity, the type of security on which it rests, and the greater formality with which it is issued. The notes of a corporation, apart from the promissory notes of a few months' tenor, partake of most of the characteristics of the bond. The majority of them however run for a few years only, whereas bonds may not mature for as many as forty years after their issue; also the notes frequently pay a higher rate of interest than do the long-time bonds. This arises from the fact that note issues are resorted to at a time when the ruling rate of interest is too high to make it desirable to issue bonds; by the time the notes mature it is expected that the interest rate will have fallen so that, if necessary, long-time bonds may be sold to replace them.

The bondholder has no voice in the management of the business so long as his interest payments are met with regularity, except in so far as his contract with the corporation may limit its freedom of action in respect of matters which have a bearing on its ability to pay the principal or interest of the bond. If the security behind the bonds is a mortgage on the physical equipment of the concern, as is frequently true, default in the payment of either interest or principal may force the bondholders to foreclose the mortgage and take over the operation of the business to protect their own interests. In case of the liquidation of a business the bonds are redeemed before either preferred or common stockholders receive anything.

Of these three groups of securities the bond gives the least risk of failure to receive payment of income and amount invested, the smallest return, and no control. The preferred stock carries with it more such risk, a greater return, and usually no control. The common stockholder has the greatest such risk, the possibility of the highest return, and usually complete control.

The property account of a corporation. Many of the features of fundamental importance in connection with the corporation can be illustrated to best advantage by the use of a simple property account, or balance sheet. A property account is a record, as of a particular time, of a real or fictitious person's *assets* and *liabilities* and of the relationship between them. Assets consist of wealth and property rights against other free persons and their wealth; a person's liability is a property right of another against him or his wealth. Thus if A owns a house, it appears on the asset side of his account; his promissory note to B for $500 appears on the liability side of his account and on the asset side of B's account.

Referring to our former example we will assume that the Steel Manufacturing Company has been incorporated with a capital stock of $1,000,000, all of which has been paid for in cash. The property account would then appear as follows:

STATEMENT OF STEEL MANUFACTURING CORPORATION
JULY 27, 1915

Assets		*Liabilities*	
Cash	$1,000,000	Capital Stock	$1,000,000

Capital stock appears as a liability since the corporation is considered as a distinct entity, and its capital stock represents rights which the owners have against this artificial person. The cash appears as an asset, since the corporation owns it.

Now let us assume that the corporation markets $500,000 worth of its bonds, and spends $1,200,000 for plant, including machinery, etc. Its balance sheet would then read:

STATEMENT OF STEEL MANUFACTURING CORPORATION
JULY 31, 1915

Assets		*Liabilities*	
Plant	$1,200,000	Bonds	$ 500,000
Cash	300,000	Capital Stock	1,000,000
	$1,500,000		$1,500,000

We have now two types of liabilities — the capital stock, which is a property right of the owners, and the bonds, which are property rights of outsiders.

During the next ten years the plant is in operation making and selling the goods for which it was built. It has sold some of its goods and has purchased some of its raw materials on credit, it has on hand a supply of materials of various kinds, raw and finished, and it has increased its investment in the plant. Furthermore the years have been profitable, and we find this reflected in the balance sheet of July 31, 1925.

STATEMENT OF STEEL MANUFACTURING CORPORATION
JULY 31, 1925

Assets		*Liabilities*	
Plant	$1,800,000	Bonds	$ 500,000
Raw Materials	100,000	Interest due Bondholders	30,000
Finished Goods	150,000	Accounts Payable	60,000
Accounts Receivable	75,000	Notes Payable	50,000
Notes Receivable	25,000	Capital Stock	1,000,000
Cash	102,560	Surplus	500,000
		Undivided Profits	112,560
	$2,252,560		$2,252,560

Many of the items on this balance sheet need explanation. The items " accounts receivable " and " notes receivable " among the assets are property rights, arising probably from the sale of goods, which the corporation has against other persons and which are evidenced respectively by entries on the books of the corporation and by promissory notes. The corresponding items among the liabilities refer to similar obligations to others on the part of the corporation. The other two items which need explanation are surplus and undivided profits. It will be noted that our assets total $2,252,560, and that there are claims against the assets on the part of outsiders of $640,000 (bonds, interest due, accounts and notes payable) and on the part of the shareholders of $1,000,000. If we subtract the sum of these amounts ($1,640,000) from the total assets we have left $612,560, which must evidently represent the net increase in the value of the assets since our last statement.

This increase in the value of the assets belongs to the shareholders. The corporation is a fictitious person and as such can never have assets greater than its liabilities; they must always balance. To indicate the growth in the value of the stockowners' holdings we might possibly add the $612,560 to the capital stock, but it is desirable to leave this item intact, in part to show the amount of the original investment, and partly for technical reasons. These two new items are therefore introduced to represent the growth in the owners' share in the assets of the corporation.

The only distinction between surplus and undivided profits is that which is dictated by convenience. Surplus is kept at a round figure and usually represents a reinvestment in the business; undivided profits are not definitely committed to any use but are generally considered as available for the payment of dividends, and they vary as profits are earned and paid out or losses suffered.

The capital stock, surplus, and undivided profits added together represent the *owners' equity* in the corporation, or the *net worth* of the corporation to the owners.[1] The net worth may always be found by subtracting the liabilities to all others than the owners from the total of the assets. The reader will observe that, in any property account or balance sheet, the asset sides is a list of property rights against wealth and persons, whereas the liability side shows to whom these property rights as a whole belong. The liability side is thus a record of *proprietorship.*

---

[1] The word "capital" is a much abused term, used in various senses which have become so entrenched in popular and even scientific language that we unfortunately

Losses and insolvency. The relationship between these items representing ownership may perhaps be made more vivid by taking a case of declining assets. Referring again to the last statement on page 105, let us suppose that a note for $10,000 proves to be uncollectible, making it necessary to reduce notes receivable and the total assets by $10,000. Losses fall first upon the owners of the business, and we will therefore reduce the undivided profits by $10,000 to make the two sides balance. A fire occurs and wipes out the plant entirely, causing a loss of $500,000 not covered by insurance. Again the total assets will be decreased by $500,000, the $102,560 still remaining in undivided profits will be wiped out, and surplus will be reduced by $397,440, thus balancing the account.

Losses can continue to occur without affecting the outside creditors' position until the point is reached where capital stock, surplus, and undivided profits have been wiped out entirely. When that point is passed the corporation is said to be insolvent. In precise terms a corporation is insolvent when its assets are less than its liabilities to others than the owners. When the assets are greater than the liabilities to outsiders the corporation is solvent; when they are just equal to the liabilities to the outsiders, the corporation is said to be on the margin of insolvency.

The value of stock. It has been customary to assign a definite value to each share of stock, and to engrave this value on the stock

---

cannot restrict it to one meaning without too violent a break with common usage. We have already defined capital as one of the factors of production (See Chapter II). We must now note that the term capital may be used also to designate what is in the text called the "net worth" of a corporation or of a partnership or individual. This should always be distinguished from the "capital stock" when it is not the same as the net worth. And finally we may use the same overworked term to designate all the resources at the disposal of a business enterprise, incorporated or not; that is, the total assets.

certificate. This value is known as the *par value*, and the par value, multiplied by the number of shares outstanding, gives the figure at which the capital stock is carried on the books of the corporation. While there are certain conveniences which attach to this custom, it has tended to confuse certain investors and others as to the true worth of a share of stock as reflected in either its actual or potential earning capacity. To engrave on a certificate of stock the words "Shares, $100 each" may be but the expression of a pious hope; the shares may be worth no more than the paper on which they are printed, or they may be worth much more than $100 per share.

A more reliable index of real worth than par value is found in the *book value*, which gives the accountant's view of the value of each share of stock. This is easily obtained by taking the value of the owners' equity and dividing it by the number of shares. According to the last balance sheet above the owners' equity was $1,612,560, and there were 10,000 shares outstanding; each share must then have a book value of $161.256. The book value fluctuates with the increase or decrease in the owners' equity; large earnings swell surplus and undivided profits and send the book value up, while losses, which of course must be borne by the owners, decrease the surplus and undivided profits and cause book value to decrease.

The book value will represent the real value of stock to the extent that the accountant's estimate of the value of the assets is correct. In some items there can be no possibility of error; if there is cash on hand or deposits in the banks the accountant has no latitude. But in determining the figures at which the plant shall be carried in the statement and in estimating the proportion of bills and notes receivable which are uncollectible there is room for considerable error. The conservative accountant has a tendency to undervalue the assets in order to be on the safe side; he may carry the real estate at the same figure for ten years or more, even though real estate values in general have increased greatly. On the other hand the fly-by-night concern may greatly overvalue the assets in the attempt to bolster up the selling price of its stock, or to improve its credit standing, or possibly to conceal transactions which have " milked " the corporation for the benefit of insiders. The book value of a share of stock, though in general much more significant than the par value, may be considerably greater or much less than the real value.

A third estimate of the value of a share of stock is its *market value,* the selling price of the share in the open market. This is based on the estimate of the dealers in the market of the earning power of the corporation. If there is an active market for the stock of a particular corporation, enough interest will be manifested by investors and professional traders in the affairs of the corporation so that the market value, aside from temporary speculative fluctuation in either direction, may be taken as a fairly trustworthy indication of its true value.

Interest payments and dividends. We have already stated that the payment of interest on bonds is a part of the contractual relation between the corporation and the bondholder. If we assume the interest rate in our example above to be six per cent per annum, payable annually, the corporation evidently has a yearly charge of $30,000 to meet. The payment of $30,000 for this purpose would change our balance sheet as follows: (1) cash would be decreased by $30,000, and (2) the item " interest due bondholders " would be wiped out entirely.[1] A simultaneous and equal change in both assets and liabilities leaves the totals equal, and no further change in the statement is required. Let us assume further that the directors decide to pay a five per cent dividend to the stockholders. This means five per cent on the capital stock, which is $50,000, or $5 per share of stock. Cash would be decreased by $50,000, and undivided profits would be similarly decreased. The owners' equity is reduced to $1,562,560, and if we calculate the book value of each share of stock we will see that it is now $156.256 (*i.e.,* $1,562,560 divided by 10,000); in other words it has decreased by precisely the amount paid in dividends on each share of stock.

---

[1] From the fact that the item "interest due bondholders" in the statement of July 31, 1925, is $30,000, the exact amount of the annual interest charge, we infer that this is the date on which the annual payment of interest is due. On any other date this item will be something less. Immediately after the payment, it is zero, as stated in the text; thereafter it increases gradually, as the liability "accrues," until it again reaches $30,000 on the date when payment is due one year later.

Stock dividends. Suppose the Steel Manufacturing Corporation decides to declare a *stock dividend* of fifty per cent. Each shareholder will receive half as many shares as he now possesses. A, who owns 5,000 shares, will receive 2,500, and the other shareholders in proportion. It is quite evident that this does not disturb the assets in the slightest; so far as our statement is concerned all the changes will be among the liabilities. Capital stock will be increased to $1,500,000 and the surplus of $500,000 will be wiped out entirely. The net worth of the business is not altered, but the book value of each share of stock is decreased, for now we divide the net worth by 15,000 instead of by 10,000.

It is quite evident that the position of the stockholders has not changed at all so far as their claims upon the assets of the corporation are concerned. The 5,000 shares which A owned had a book value before the stock dividend of $161.256 per share, or a total value of $806,280; now he holds 7,500 shares, each having a book value of $107.504, or a total value of $806,280. The book value per share of stock has declined, but the total book value is not affected. In fact the stock dividend, except in a figurative sense, is not a dividend at all, but merely a transfer from surplus to capital stock.

Since the stock dividend makes no change in the position of the stockholder with respect to the corporation, it may be asked what purpose it can serve. Without attempting a complete answer, the two most common purposes may be cited. In the first place it may happen that the market value of the stock is so far in advance of the par value that a share of stock is less easily traded in. Five shares which sell for $100 each are much more convenient than one share which sells for $500. The stock dividend may then be resorted to in order to increase the number of the shares and lower the value of each. But the chief motive is probably to avoid the appearance of unusually high profits and high dividend rates. A dividend of 30 per cent on a capital stock of $1,000,000 is no greater than a dividend of 15 per cent on a capital stock of $2,000,000, but it seems greater; the former may attract considerable notoriety, lead to the suspicion of monopoly, and invite competition, while the other passes by without exciting comment.

Watered stock. Stock is said to be "watered" when the corporation issuing it has not received full value for it. There are several different ways of watering the stock. One of the most common is to issue stock in payment for property or services which are overvalued. This may be done innocently and in good faith, or it may be done with intent to defraud. Suppose that $100,000 worth of stock, par value $100 per share, is sold to investors for cash, and that another block of $100,000 is given in payment for a tract of land which, at the most liberal estimate, is worth not more than $50,000. The statement would show assets of $200,000 and liabilities of $200,000, and the book value of each share would be $100. A true record on the books of the corporation of the value of the assets would however record them as worth $150,000 and would thus show a book value of only $75 per share. A stock dividend without a sufficient surplus to justify it is a form of stock watering. In our example above a stock dividend of a million dollars could have been declared only by raising the recorded value of the assets by $387,440.

The essence of stock watering is overvaluation of the assets of the corporation, and it may be resorted to for any of the purposes stated in a previous paragraph dealing with book value. Like the legitimate stock dividend, its purpose may be to conceal large profits, though in this case it really conceals them instead of merely avoiding an appearance of high profits. Whatever its purpose, stock watering involves falsification of accounts and carries at least the presumption of dishonest intent. Stock watering is always, and particularly in the case of a public utility such as a railroad or a gas company which may receive an exclusive grant to operate in a particular territory, against the public interest. The public has the right to know the true value of the corporation's assets and the rate of its profits and whether the prices paid for its services are enabling the company to earn exorbitant profits.

The corporation as a factor in modern economic life. The utility of the corporate form in modern business manifests itself chiefly through its influence in directing the savings of multitudes of people into the hands of business men. The limitation of liability in the case of insolvency has made investment in corporate stocks very attractive and has been an important means of gathering in savings from diverse sources. Formerly an investor had to assume risks all out of proportion to his knowledge of the particular business, and hoarding was often the only practical way open to the small investor to conserve his wealth for future use. Today he can invest directly in corporate securities or indirectly through the medium of savings banks and insurance companies.

The small denominations in which shares of stock and bonds are issued and the ease with which they may be transferred facilitate investment in corporate securities. One or two or three men may shun an unproved but attractive venture, whereas a thousand men acting together may be drawn to it, for each one knows that his stake in the enterprise will not cripple him financially even if it be entirely lost. He knows further that as a rule he may sell his interest at any time he chooses to one whose faith is greater and thus withdraw before suffering the total loss of his investment. Advance in industry and commerce is conditioned upon the taking of risk, and any device which so divides and scatters the risk as to make it less burdensome to the individual will in the long run be an aid to production.

The ease with which a corporation engaged in a legitimate line of business can secure capital leads to large scale production, the further division of labor, greater specialization in production, and therefore to the possibility of a greater flow of goods at a lower labor cost. We think of large scale production as the result of the Industrial Revolution and the introduction of factory methods of production, and it is true that large scale production is possible without the corporate form of organization, but it must be admitted that the corporation has been a factor of no small significance in its development.

Although the corporate form of organization has stimulated capital accumulation and has promoted the economical use of the resources of society, it has also created special problems of some importance. The delegation of the management of a business may not result in the most efficient management. An individual running his own business and with his own funds invested in it, has the most effective stimulus to try to secure as large a profit as possible. In a corporation the management may be satisfied with modest profits — enough to pay the customary dividends on the common and preferred stocks — and therefore not strive to attain those economies which might make larger profits obtainable.

The corporation, with its transferable shares of stock, facilitates the concentration of financial power in the hands of a single individual or of a group of individuals and therefore may lead to the use of financial power for the oppression of competitors and the attainment of a position of monopoly. It is indeed difficult to see how a monopolistic control in any large industry could be built up on the basis of individual proprietorships or partnerships. The expense of securing the controlling interest in partnerships and the difficulties which would ensue on the death of the individual proprietor or a partner seem almost insuperable to the attainment of this object. In contrast with this there is the ease of transferring shares of stock in the corporation, the possibility of obtaining control of the corporation by the control of a part of the outstanding voting stock, and the longer life of the corporation. One is led to the conclusion that the monopolies of the modern era have depended in large measure on the corporate form of organization and is not surprised to find that the monopolies which the United States Government has seen fit to regulate or suppress have utilized the corporate form in one way or another.

The marketing of corporate securities. The social service which corporations render in facilitating investment would be seriously impaired if they did not have the assistance of a multitude of specialized banking houses, investment brokers, and other middlemen between them and the investing public.

The services of the middleman are not always required. Many a corporation has a very small capitalization and does not try to interest the public in its stock; the small group of men who have organized the corporation and who expect to become its responsible officers subscribe for the whole issue. Other companies are local in character and are able to appeal to local investors without the assistance of middlemen. Even in some large concerns the ownership of the stock is concentrated in a few hands; the stock of the Ford Motor Company is held entirely by the members of the Ford family. Yet in most large corporations there is a wide distribution of ownership — there are for example more than 165,000 owners of shares of stock in the United States Steel Corporation — and a company with a capitalization of several millions is frequently too large to be financed either by a group of individuals or by a small community. An appeal must be made to a wider public, and bond houses and investment bankers, as well as the brokerage houses and the stock exchanges, exist to carry that appeal to the investing public.

Selling stocks and bonds is a type of merchandising, a highly specialized type, in which a detailed knowledge of the goods to be sold and a wide acquaintance among buyers are prerequisites for success. The corporation has neither the organization for selling its securities, nor the knowledge of the market which is desirable. True, it could install a security sales department, but the occasions when it has securities to sell are rare; *i.e.*, when the corporation is formed or when new issues are floated to secure additional capital. The financial middlemen are constantly dealing in a wide range of securities, are in connection with investors all over the country, and can handle the whole business more cheaply than the corporation itself. It may take the corporation a longer time to sell its securities, and, since it presumably needs the funds at once, this may decide it to put the matter in the hands of bankers who will agree to hand over a certain sum of money by a certain date. Furthermore if the corporation experiences any difficulty in disposing of its securities, it may acquire a bad name which will militate against success in marketing future issues.

Underwriting. We have not the space, nor is it properly a function of an elementary text in economics, to discuss in detail the types of investment houses and the special functions of each type. We must treat the subject in a very broad way and shall use the terms " investment houses " and " investment bankers " in a loose way to include all the houses which assist in selling securities to the public. Let us understand clearly however that the investment houses do not themselves invest in the stocks and bonds which they carry unless circumstances compel it. They are middlemen, intervening between the corporation in need of funds and the investing public.

Having this in mind we can proceed to a brief discussion of the way a large issue of bonds might be handled. Let us assume that a corporation has submitted a report to a prominent banking house which specializes in financing investments in industry and has satisfied it that the purpose for which the proceeds of the bond sale are to be used is sound and likely to be profitable. There are several million dollars of bonds to be sold, and the banking house is unwilling to assume all the responsibility for their sale. It therefore organizes a syndicate or group of investment bankers, located in the same city or in other cities, which agrees to underwrite the sale of the bonds.

"To underwrite is to insure, and, as the term is used in the financial world, *underwriting may be defined as insuring the sale of a corporation's securities in advance of a public offering.*"[1] Practically the syndicate agrees to take the issue of bonds and make payment for it at a specified time and at a fixed price per bond. Naturally it expects to sell the bonds before the date when payment is due, but if it fails to do so or if it fails to sell the bonds at the price anticipated the loss falls on the syndicate and not on the corporation. The assumption of this risk is the essential feature in the underwriting agreement.

The syndicate expects in the natural course of events to make a profit from handling the bond issue, and if it agrees to pay the corporation ninety-five dollars per bond, it may fix upon ninety-eight dollars as its selling price to the public, thus leaving it a commission of, let us assume, one dollar for selling and two dollars for underwriting the sale of each bond.

Each member of the syndicate agrees to be responsible for the sale of a given proportion of the total bond issue and therefore has an incentive to bring into play all of its own sales organization and to interest the smaller houses with which it has dealings. Advertising and personal solicitation by salesmen play their part here just as they do in selling a particular brand of coffee. Naturally the wider the appeal which is made to the public the greater is the chance to dispose of the bonds before the date when payment has to be made to the issuing corporation. Our financial organization is so closely tied together that for the sale of a very large issue of bonds bond houses and investment bankers in almost every city of the United States can be interested.

---

[1] L. H. Haney, *Business Organization and Combination*, p. 301.

The investment houses act as middlemen; they do not expect to tie up their own capital in these securities, but sometimes that is unavoidable. If the issue has not been sold before the day arrives when the corporation must be paid, the syndicate must make payment in full, either from its own resources or by borrowing. The latter is the more common method, for by pledging the unsold bonds as collateral with the large banks it can borrow up to eighty or ninety per cent of their sale value. At the same time the sale of bonds goes on, and as funds are received they are used to reduce the amount of the loan.

Other marketing agencies and other functions. There are other marketing agencies than the ones which we have mentioned which serve to bring the leader of industry and the saver together, and other means than the ones we have mentioned which may be used for this purpose. Our purpose has been to point out the place which the financial middleman occupies in our economic organization, rather than to describe the financial organization in detail.

It should, perhaps, be noted in passing that the organization which we have been describing serves as well for the buying and selling of securities already issued as for the sale of new securities. But the sale of old securities has no interest for us at this point, except in so far as the facilities for their exchange assist in the sale of the new ones. Securities issued in the past have absorbed the savings of the past, but the current issues of securities must be adjusted to current savings.

The stock exchanges. A discussion of the marketing of securities cannot be considered adequate without mention of the facilities offered by the stock exchanges for the quick disposal of stocks and bonds of the better grades.

The New York Stock Exchange, which we shall use for purposes of illustration, dates back to 1792, when a few brokers made an agreement governing their relations with their customers. It is a voluntary association of 1,100 members who conduct business according to rules and regulations adopted partly in conformity to legislation and partly to preserve a high code of business ethics. A Governing Committee is clothed with power to see that this code is adhered to, and for serious infractions of it a member may be expelled and his "seat" vacated. No one not a member of the Stock Exchange is allowed to do business on the floor of the Exchange. Most of the members are brokers and buy and sell at the order of a principal, although there are some who trade independently, buying and selling on their own account.

A stock or bond must first be *listed* on the Exchange before it can be dealt in on the floor. This involves an application by the officers of the corporation concerned, which is referred to a committee of investigation, according to whose report the application is either approved or denied. A stock which is listed on the Stock Exchange seems to the public to bear a certain stamp of approval, although the Stock Exchange intends (presumably) to do no more than credit the management with honest intentions in the sale and disposition of its stock. It never guarantees the worth of the stock or bond so listed or recommends its purchase. The Illinois Securities Law of 1919 quite aptly summed up the significance to be attached to listing, when it classified securities which are dealt in on the New York and certain other exchanges in that group of securities whose inherent qualities ensure their sale and disposition without fraud.

The stock exchange does not float new issues of stocks and bonds; these have already been sold before they are listed on the exchange. A stock exchange serves as a place where those already issued may be bought and sold. Among other advantages which a stock exchange affords the following are of particular importance in our study. In the first place it provides a market where any quantities of listed stocks and bonds may be sold almost immediately for cash. This serves to make investment in securities attractive, for the investor knows that in case of need his securities can be sold at once or can serve as collateral for bank loans if he does not choose to sell them, since bankers know they can sell them at any time it is necessary to protect themselves. And secondly the prices of stocks as determined on the exchanges aid in the direction of capital investment. If the price of a stock is consistently high it means that profits in that industry are above the average. When therefore a new company is being launched, it has a better chance to secure the capital it needs through the sale of its securities. On the other hand consistently low prices of a stock may indicate an overinvestment in a particular line, and those having funds to invest avoid it. Thus the most profitable and probably the most economical investment of savings is promoted.

The services of investment bankers. Directly or indirectly a great many different individuals are adding each year to the total amount of capital which producers and others are using. The investment bankers act as the middlemen in directing this capital to the proper enterpriser. The world " direct " is not used unadvisedly, for the bankers do not play a passive rôle by any means. The risks inherent in underwriting make it necessary for them to scrutinize each proposition carefully before committing themselves to it. Their wider information regarding investment possibilities and their more profound knowledge of finance tend to prevent misdirection and dissipation of capital. By refusing to underwrite securities for corporations in a line which is already overdeveloped or by refusing to handle the issue of a corporation in whose officers they have no confidence, they direct the flow of capital and throw the control of industry into the hands of men best equipped to manage it. Doubtless mistakes are made, but they are fewer than would occur were this function assumed by the corporations themselves. Bankers are impelled by self-interest to avoid losses and to safeguard their reputations, to take an impartial view of the situation as a whole, and to use their best judgment in directing the flow of savings.

The investment banks and the allied agencies are an essential part of our coöperative system and represent one phase of the application of the principle of division of labor. They have been called into being to assume burdens which were becoming too heavy for the corporation and they perform them in a more satisfactory manner and more cheaply than could the corporation. It is questionable whether the astonishing developments in production during the nineteenth century could have come about without the services of the banks in bringing large and small savings together and making them available for investment in industry.

[illegible] out

EXERCISES

Hand in Wed

1. Arrange the following items in the form of a statement of the assets and liabilities of a corporation: land and buildings, $90,000; capital stock, $100,000; bonds, $50,000; machinery and equipment, $75,000; Liberty bonds, $5,000; cash, $7,000; surplus, $60,000; taxes due, $3,260; accounts receivable, $12,000; raw materials on hand, $26,000; finished goods on hand, $20,000; accounts payable, $15,000; undivided profits, $6,740.

2. There are 1,000 shares of capital stock outstanding. What is the par value of each share; the book value? *Note:* In this and each of the following exercises, refer to the original statement as in Exercise 1.

3. A cash dividend of six per cent is declared and paid to the stockholders. Make the necessary changes in the statement. What is the book value per share of stock now?

4. The company declares a stock dividend of fifty per cent. Make the necessary changes in the statement. What is the new book value per share? Stockholder A held ten shares of capital stock before the stock dividend was declared. Calculate the book value of his holdings of capital stock before and after the stock dividend.

5. A fire breaks out and destroys the stock of raw materials and the warehouse, valued at $5,000, in which they were stored. These losses are not covered by insurance. Make the necessary changes in the statement. Is the company still solvent?

6. Make such changes in the balance sheet as would be necessary to illustrate stock watering.

CHAPTER VII

LARGE SCALE PRODUCTION AND COMBINATION

The growth of large scale production. A consequence of the Industrial Revolution has been a growth in the size of the producing unit as measured by the amount of capital used, the number of workers employed, and the volume of output. The table on the opposite page gives a statistical view of this growth for cotton manufacturing and for the boot and shoe industry in the United States for recent periods of fifty and thirty years respectively.[1]

A study of these statistics reveals several interesting features. There was a positive decrease in the number of establishments devoted to cotton manufacturing in the period from 1859 to 1899, and for the period as a whole there was not a great change in the number of establishments — an increase of only about 20 per cent. During the same half century the average number of workers per establishment increased by over 150 per cent, the average amount of capital per establishment increased by about 660 per cent, and the average value added to the product by manufacturing per establishment increased by more than 300 per cent.

In the boot and shoe industry the situation is nearly parallel. In the generation following 1879 the number of workers per establishment increased by 140 per cent, the average amount of capital per establishment by nearly 600 per cent, and the average value added to the product by manufacturing per establishment by 300 per cent. The number of establishments in the industry shows a significant decline, whereas for cotton manufacturing there is, for the same period (1879–1909), a great increase. Part of the explanation of this difference is found in the fact that the cotton industry was on a factory basis at the opening of the second half of the nineteenth century, whereas shoe manufacturing was still in large measure a local industry carried on in a very small way. The application of factory methods to shoe manufacturing meant the extinction of a large number of small producing units by a smaller number of units operating on a larger scale.

---

[1] Abstract of the Census of Manufactures, U. S. Bureau of the Census, 1919. Derived from tables on pp. 52 and 591. Values have been reduced to a common unit by using the index numbers of the Bureau of Labor Statistics (Bulletin Number 181, p. 266), base year 1914.

The figures given for the average amount of capital are admitted by the Bureau of the Census to be very defective, but they are the best available, and are therefore included.

Reports of the U. S. Census are the source of the statistics presented in the pages immediately following.

Year	*Number of establishments**	*Average number of wage earners per establishment*	*Average amount of capital per establishment*	*Average value added to product by manufacturing per establishment*
		Cotton Manufactures		
1859	1091	112	$ 86,059	$ 50,976
1869	956	142	88,664	41,435
1879	756	228	278,333	120,095
1889	905	262	460,214	146,986
1899	1055	287	582,738	202,721
1909	1325	286	653,710	205,682
		Boots and Shoes (Leather)		
1879	1957	57	$ 22,191	$ 32,830
1889	2082	64	53,840	57,559
1899	1599	89	82,138	74,336
1909	1343	138	154,480	129,453

* As a rule the term "establishment" represents a single plant or factory, but in some cases it represents two or more plants which were operating under a common ownership.

We find this same transformation taking place in a number of other industries. In the period from 1849 to 1919 the number of establishments making agricultural implements declined from 1,333 to 521, the number of cooperage establishments from 2,902 to 1,099, and the number of iron and steel blast furnaces from 404 to 195; but in every instance the industry shows an increase in the number of wage earners employed, in the amount of capital invested, and in the value of the output.

These figures show pretty clearly the increase in the average size of the industrial plant, but they do not demonstrate the significance of the very large plant as compared with the plant of small or moderate size. This is brought out by the table given below.

ESTABLISHMENT GROUPED BY VALUE OF PRODUCT
ALL INDUSTRIES. 1919

Value of product	*Percentage of total number of establishments*	*Percentage of value of product to total*
$5,000 or less	22.6	0.3
5,000 to $ 20,000	30.1	1.5
20,000 to 100,000	26.9	5.7
100,000 to 500,000	13.7	14.4
500,000 to 1,000,000	3.2	10.4
1,000,000 or over	3.6	67.8

The establishment of small size is seen to be most numerous — those producing $20,000 of goods or less a year formed nearly 53 per cent of the total number of productive units, while those with an annual output of over a million dollars constituted only 3.6 per cent of the total. But the plants with a yearly production of goods valued at $20,000 or less contributed only 1.8 per cent of all the goods produced, whereas the plants with an output of $1,000,000 or more were responsible for the production of nearly 68 per cent of all the goods produced, measured in terms of value.

The picture may be made more vivid if we approach it from a still different angle. In 1919 the establishments employing from one to five laborers comprised 49 per cent of the total number, whereas the establishments employing over 500 workers formed one per cent of the total, but these establishments employed 40 per cent of the total wage earners in industrial plants as contrasted with the 3.4 per cent employed by the establishments in the former group. Establishments employing 1,000 laborers or more formed only four-tenths of one per cent of the total number of establishments, but they employed 26 per cent of the total number of workers.

Two conclusions of considerable importance can be drawn from these figures. The first is that small scale industry is still of significance in our national life, particularly to supply local markets. "Big business" has not entirely supplanted the small producer even in the lines of industry where the advantages of large scale production seem most marked. In steel works and rolling mills, where of all the industries of the United States the largest percentage of large scale establishments is found, there were in 1919 but 202 establishments out of a total of 500 which employed an average working force of 500 or more, and in the cotton goods industry only 240 out of a total of 1,496 establishments employed five hundred or more workers.

At the same time it seems true that the large unit has been growing, both absolutely and relatively, at the expense of the smaller-sized units. Production for national markets has been supplanting production for local or regional markets, and with this change has come the decline of the small unit and the substitution of the larger.

Conditions necessary for large scale production. One of the first prerequisites for large scale production is obviously a product which is in great demand, for unless a large output can be marketed profitably there is no scope for increased refinements in the productive apparatus.

Another requirement is that the article be not one which is demanded because of the finish which hand labor may give it. It must be amenable to the machine process, capable of standardization, and one which does not lose attractiveness in the eyes of the purchaser thereby. Many articles of clothing, of personal adornment, and of luxury can never be produced by large scale methods.

Economies of large scale production. Large scale production, where it is found, exists for the simple reason that it is more efficient than small scale production, and by this we mean that economies in production are effected by the larger unit which make it possible for the good to be sold at a lower price. This not only extends the range of commodities over which the consumer has command and therefore enlarges his living, but it is advantageous in a more profound sense, for it implies the more effective utilization of our limited supplies of labor and capital.

Power and machinery. One great economy of large scale production results from saving in the use of power-driven machinery.

There are savings in the use of power, as well as in the use of machinery. The saving in power arises from the lower unit cost of installing and operating a power plant. It is a matter of common knowledge that the ordinary heating outfit for a dwelling house is relatively uneconomical. A central plant for a group of closely associated buildings costs more to install, though probably not so much as the combined cost of heaters for the separate buildings, and when it is installed it is much more economical to run; per ton of coal consumed more heat is applied to its proper use than in the former case. The same holds true of the large factory. The plant costs more to install than for the small factory, though not proportionately more, and greater efficiency in the production of steam results. Hence the proportion of the total cost of power which is to be attributed to each finished article is less.

Machinery can be used for more of the processes in a large factory than in a small one, for if a machine is to be used to advantage it must be used continuously. In determining the unit cost of production by machine one must take into account the yearly interest on the investment and also a sum large enough to cover repairs and the eventual replacement of the machine, as well as the power and labor necessary to run it. If the machine is used only a part of the time the proportion of the cost which must be borne by each unit of the good undergoing the particular process may make its use more expensive than hand labor. A large factory is able to employ machines which save only a fraction of a cent over the cost of hand labor but which are distinctly worth while on account of the large volume of production. This is one way of stating the proposition that large scale production permits of the greater application of the principle of division of labor. There can be greater specialization by workers, and processes can be more finely divided and therefore more easily subjected to machine methods.

By-products. Another economy lies in the utilization of by-products. The meat packing industry offers one of the most illuminating examples of this. Swift & Company give the following summary of the products derived from a steer:

Percentage of Weight of Finished Products to Weight of Live Steer

Beef	54.3%
By-products:	
Hide	5.9
Fats	2.2
Head	2.2
Feet	1.1
Blood	.7
Casings	.8
Miscellaneous	3.2
Valueless	10.1
Shrinkage	19.5
	100.0

A list of the finished goods which are made in whole or in part by Swift & Company from these by-products includes the following: glue, fertilizer, oleomargarine, lard compounds, soaps, glycerine, leather dressing, animal feed, buttons, combs, gelatine, ice cream, candy, pharmaceutical products, snuff containers, gold beaters' skins, and sausage containers. In the early days the materials from which these goods are made had to be carted from the city and buried; so far from being a gain, they were the occasion of considerable expense.

It is reported that the Ford Motor Company has developed a business in by-products which, during 1924, brought in more than $13,000,000. Creosote, methyl acetone, acetate of lime, benzol, and ammonium sulphate are among the products mentioned. The sale of the surplus factory scrap metal and non-metallic material is said to reach $4,000,000 a year.

Only when the quantity of by-products is large is it worth while for the company concerned to refine them. Otherwise they are thrown out as waste or perhaps sold to a subsidiary industry which has started up to handle the by-products of a number of similar industries. The large industry however is likely to make a greater profit from refining its by-products itself than the small producer can get by selling them to another concern. Unless competition to secure the by-products sets in, the small producer finds it necessary to get rid of them as best he can, and he is usually glad to take any price offered in order to avoid the expense of disposing of them as waste.

The economy which results is plain; any profit which is obtained from the by-products diminishes by so much the price which must be obtained for the principal product in order to cover the expenses of production. The meat packers for example often pay more for live cattle and hogs than they receive for the dressed meat, but they make up the difference in the return from the by-products.

Technical experiment and research. A large company is better able to conduct experiments than a small company. Experimentation is a risky and sometimes an expensive affair. Many trials are necessary before a device or a process is developed to the point where it is of practical use. For a small concern the financial burden may be too heavy to carry, whereas the company with greater resources can carry the load through the lean years, knowing that in the long run it will more than pay for itself.

In the making of pig iron chemists were not employed in this country to analyze the ore and the materials used with it in smelting until after 1870. The blast-furnace manager was supposed to be able to diagnose the condition of his furnaces by instinct. The result was that inferior grades of ore oftentimes enjoyed a better reputation than those of better quality. A foreman who had evolved a mixture which suited low grade ore often rejected ore with a higher iron content simply because the mixture did not work. Mr. Carnegie was one of the first to employ a chemist, and he reaped an immediate reward in that he was able to secure ores at a low price which were found on analysis to contain an unusually high percentage of pure iron. In his *Autobiography* he says that years after he had taken a chemist as his guide some of the proprietors of other furnaces claimed that they could not afford that luxury.

Another example which illustrates more clearly the expense and risk involved in experimentation is found in the shoe industry. After years of experimentation and the expenditure of more than a million dollars the pulling machine was perfected. It is stated that this machine has saved the manufacturers at least four times its cost yearly. Many other examples might be cited, but it is quite generally known that research work in the realm of pure science and in the field of marketing is growing in importance among business houses and associations of business men. The United States Chamber of Commerce has estimated that American manufacturers spend annually about $35,000,000 on work of this sort, with a probable saving of half a billion dollars in the conduct of business. The emphasis which the Department of Commerce has recently placed on the utility of still greater expenditures for research in pure science is likely to add impetus to a movement which has already attained considerable momentum.

Economy in administration. The cost of administration and supervision is not much greater for a large factory than for a small one, certainly not proportionally greater. Within limits the head of a department can generally take care of a large department with about the same effectiveness as of a small department. The chief additional expense consists in a larger clerical force. Similarly the cost of buying and selling raw materials and finished goods may be decreased per unit. Buying in large quantities means a real saving to the seller in packing, shipping, and sales costs, and he can therefore afford to quote a lower price to one who buys on a large scale. In selling the finished good there is better opportunity to coördinate the activities of salesmen, thus covering a given territory without overlapping and with greater effectiveness, while national advertising can be used with greater profit.

An interesting statistical study of the costs of production of large and small plants is presented in the report of the Federal Trade Commission on the canned salmon industry. In 1916 the unit cost of production for large plants was $3.491, and for small plants $4.111; in 1917, the figures were $4.302 and $5.681 for the large and small plants respectively.

The field of large scale production and its limits. Our attention thus far has been devoted exclusively to manufacturing, the field in which large scale production obviously finds its greatest opportunities and from which may be drawn the most illuminating illustrations of its principles. Many other lines of industry have however proved themselves amenable to large scale production. In transportation, banking, and marketing the size of the organization has greatly increased in the last century. Doubtless this increase in size has come partly in response to the growth in the size of the manufacturing establishment and is partly a cause of that growth. In any event concentration of manufacturing in large plants requires more effective service than could be afforded by the transportation, banking, or marketing system of a century ago.

The one great industry which has never proved a profitable field for the application of large scale methods is agriculture. There are of course some examples of farms where hundreds of workers are employed, but they are the exception and not the rule. The difficulties of supervision on the farm are so great that unless the laborer is working with or near the owner he is tempted to loaf on the job. In a factory individual supervision by the foremen and automatic checks such as time clocks, machinery with controlled speed, and the piece-rate system of wage payments do away with this difficulty to a great extent, but on the farm the spur of individual gain seems necessary to secure continuous effort. Hence we find that in most cases which do approach large scale farming the tenant system is used. The owner of the farm provides materials and tools, and the farm laborer receives a certain proportion of the output, thus giving him a direct incentive to increase his efforts.

Large scale production has its limits, even in the fields most favorable to it. Probably for each industry there is a plant of optimum size, beyond which it is unwise to attempt to go. A part of this difficulty lies in the physical awkwardness of handling a large bulk of goods under one roof, but a still greater difficulty is the weakness in management which may appear. Since this weakness is common also to combinations of several plants under one management, its further discussion may better be postponed till after our investigation of the problem of industrial combination, to which we now proceed.

Combination: The problem of large scale management. A feature of modern industrial organization particularly prominent in the last few decades has been the combining of numbers of industrial units under single management. In large measure combination has been resorted to as a means of securing substantial control of the particular field of production in order to obtain monopoly profits. Two other motives however have been prominent. Of these the first has been to obtain even greater economies in production than were possible simply with large scale production; the second is to make feasible speculative profits through the issuance of large volumes of watered stock.

We have here two distinct problems. The first, and in some aspects the more important, concerns itself with the growth of monopolies — the so-called "trusts" — their effect on society, and the methods by which they may be regulated in the public interest. The second problem centres around the possible social advantages of combinations as producing units. Are there any economies in production which a number of large units can effect which cannot be obtained by the single large producing unit? If there are such economies and they are of considerable significance, this fact will have an important bearing on the first problem. It is this second topic, rather than the problem of monopolies, to which we shall give our attention in the remainder of this chapter.

Horizontal combinations. We may distinguish two types of combination, the *horizontal* and the *vertical.* As the name suggests, a horizontal combination is one which unites under a single management several plants producing the same product or products. It is a type which is quite common in the United States, being exemplified by many of the "trusts" which exist or have existed in the past, such as the American Sugar Refining Company, the United Shoe Machinery Company, the American Tobacco Company, and others.

Economy in buying. Such a combination may possibly secure a saving in the purchase of its raw materials. Buying in much larger quantities than the single plant, it may approach the producer directly and eliminate the middleman's profits, or it may secure its goods from a wider productive area and thus obtain its raw materials more cheaply. But there is certainly some limit to the savings in this line, and it is to be doubted if they are great enough to be of real consequence. Of course if the combination controls a large part of the total production it may use its power to depress the price of its raw material and cut the profits of the producers of the raw materials. But this does not constitute a saving in the social sense, as it merely gives to one class what is taken from another, nor does it give the combination an advantage over its competitors.

Economy in selling. In the sale of its products there is a possibility of saving. A reduction in the number of salesmen and in the amount of advertising and greater effectiveness in the use of both sales media may lower the sales cost. The tobacco trust, which had been formed by the combination of several large producing units, was dissolved by a court decree in 1911, whereupon the business was turned over to a number of separate companies. " The subdivision of the business provided for in the dissolution decree led to a duplication of selling organization and an increased overhead expense; and the result was a general increase in selling costs. . . . In every branch, except flat plug and Turkish cigarettes, the selling costs showed an increase. Based on the rates per thousand or per pound, the selling costs of the successor companies in 1913 in the cigarette branch were 55 per cent greater than for the trust in 1910; 91 per cent greater in the little cigar branch; 11 per cent greater in navy plug; 39 per cent greater in plug-cut smoking; 40 per cent greater in granulated smoking; 44 per cent greater in fine-cut chewing; and 46 per cent greater in snuff. The selling costs of the successor companies for flat plug and Turkish cigarettes, on the other hand, were only 76 and 83 per cent, respectively, of the costs of the trust in these branches.

" The advertising expenditures of the successor companies also greatly increased as compared with the expenses of the trust. The advertising expenses of the trust in 1910 in all branches except cigars were $10,895,132, while those of the successor companies in 1913 amounted to $23,623,564, or more than double." [1]

Economies in operation. The greatest field for economies would appear to lie in the operation of the plants. Each plant can presumably secure maximum efficiency by producing a certain number

[1] Eliot Jones, *The Trust Problem in the United States*, pp. 149–150.

of articles. If forced above that point or below it, the unit cost of goods is increased. With twelve plants operating under unified control it is possible so to arrange things that all the slack in production will fall on one or more plants, leaving the others running as usual. The American Sugar Refining Company is said to have used its Brooklyn refinery for this purpose, slowing down operations there as demand slackened and speeding up in response to increased demand. In some industries the risks and uncertainties involved in production may be reduced if there are a number of plants located in different parts of the country. A labor strike may not affect all of the plants at the same time and a tie-up of the transportation facilities of all the plants at one time is unlikely, so that if the closing of some is forced the others can be operated more intensively. In the salmon canning industry the size of the run of fish is a decisive factor, as a small run of fish may entail heavy losses. In 1917 a company operating six plants lost $115,000 on one plant, but made a net profit on the six plants of over $1,000,000. The owner of a single plant is liable to have heavy losses one year and large profits another, and the stabilization of profits which comes from enlarging the area of operations is of considerable advantage.

Specialization may be carried to greater lengths by a combination than by a large producing plant. A combination which makes a variety of products is able to specialize in factories, devoting one factory to the making of a particular article, rather than allowing all factories to make all articles, and in this way the utmost advantages of specialization are attained.

The devices or patented articles that have proved most successful can be made the common property of all the operating units, heightening the efficiency of each. With several plants it is possible to start competition between them, thus enlisting the enthusiasm of one manager to excel the others. Cost statistics may be available for all the plants, so that if one falls down in some respect the trouble can quickly be diagnosed and remedied. It is claimed that more efficient executives can be secured because of the wider field from which to choose. The interplant competition, it is asserted, will disclose men who are most capable of handling the more difficult tasks of coördination and centralization.

Saving in freight costs. The savings on cross freights are sometimes cited as proving the superiority of the combination. If under separate management there are two plants, one in New York and the other in Chicago, each may sell goods in the other's district and incur freight charges unnecessarily great. Under combination the good will be sent from the nearest point. But this saving, while probably real in many cases, can be greatly exaggerated in importance, for if the good be bulky and of low value the distance it can be shipped is limited by the cost of freight. On the other hand if it be a small and expensive object such as a watch, freight charges form such a small proportion of its value as to be of little moment.

Vertical combinations. Vertical combination involves the gathering together under one management of the various processes necessary to turn out a finished good. It is sometimes called *integration* of industry and is the opposite of that process of differentiation and specialization which has been characteristic of the factory system. In some industries there has always been a certain amount of integration. Thus in the United States spinning and weaving have commonly been undertaken by the same factory, although in England they are separate industries. In the steel industry of the United States can be found the most impressive illustrations of vertical combinations, although horizontal combinations have been by no means uncommon.

Technical advantages. There are real technical advantages to combining thus the different stages of manufacture of steel products, advantages which result in the saving of fuel for reheating the metal, of time and labor in moving or handling material, and in the more complete utilization of by-products, such as blast-furnace gas. Further advantages may be obtained through the control of sources of iron and coal. In the early history of the steel industry it was not considered essential to own iron fields, but as concentration in the ownership of the known and proved fields developed, the unpleasant prospect of being forced to pay large profits to the producers of iron ore had to be faced. The alternative was ownership or control of iron reserves. This held true of coal and coke as well. The steel industry differs from some others in that each unit not only may use the products of another as raw material but may pass its own finished or semi-finished products on to serve as raw materials for the others. Coal and iron are essential for the blast furnaces, while the ultimate products of the blast furnace are requisite in the mining and transportation of coal and iron.

The United States Steel Corporation. A picture of the extent of the process of integration can best be given by a brief summary of the scope of the operations of the United States Steel Corporation, the largest corporation in the United States, incorporated in 1901 with an authorized capital stock of $1,100,000,000. The corporation secured control of a number of other corporations by an exchange of stocks and bonds, and at the outset it controlled eleven corporations: (1) The American Bridge Company, fabricating bridge material and structural steel, (2) The American Tin Plate Company, (3) The American Steel Hoop Company, producing chiefly iron and steel bars, hoops and bands, and cotton ties, (4) The American Sheet Steel Company, engaged in the manufacture of practically a single product — sheets, (5) The American Steel and Wire Company of New Jersey, manufacturing wire nails, plain wire, barbed wire, and wire fencing, (6) The National Tube Company, producing iron and steel wrought tubing, (7) The Shelby Steel Tube Company, a consolidation of practically all the important companies making drawn or seamless tubing, (8) The Lake Superior Consolidated Iron Mines, which was an ore producing and transportation company exclusively, controlling the Duluth, Missabe & Northern Railroad (This railroad had a controlling interest in the largest fleet of Lake ore carriers, the Bessemer Steamship Company, which the United States Steel Corporation later purchased.), (9) The National Steel Company, confined largely to the production of semi-finished steel, such as billets, sheet bars, and tin-plate bars, (10) The Federal Steel Company, itself a combination controlling coal, iron ore, and transportation, and turning out about the same line of products as (11) The Carnegie Steel

Company, the most important of the group. The Carnegie Steel Company was itself the result of merger and reorganization and exemplified a high degree of integration. It had assured itself of extensive reserves of desirable ore, coal, and coke, had made a beginning in controlling the transportation of its own ore on the Great Lakes, and had acquired several railroad lines to ensure the proper and cheap handling of its raw materials and finished products. Its chief business was the manufacture of heavy steel products, such as steel rails, plates, structural steel, and crude steel which served as the primary material for other concerns, paying little attention to the manufacture of finished goods.

It is a fact of interest that almost all of the companies which were united to form the United States Steel Corporation were themselves combinations. Some of them represented merely the horizontal type of combination, such as the American Steel Hoop Company, which was a consolidation of nine concerns formed with the desire to limit competition and secure a larger profit, while some already showed signs of a fairly complete integration, as the Federal Steel Company and the Carnegie Company. The United States Steel Company was not exclusively a vertical combination, for several of the constituent companies produced the same articles.

The result of the consolidation was the creation of a giant corporation which was practically self-sufficient so far as materials were concerned and which embraced the manufacture of practically every known article in steel and iron. Its history since that time is one into which we cannot go, but we may say that in brief it consists of an extension of the processes already noted — the acquisition of more mines, furnaces, plants, steamships, railroads, and foreign selling agencies.

The question which immediately presents itself in this connection is whether the Steel Corporation has justified itself from the point of view of social saving. According to the investigation of the Bureau of Corporations into the value of the constituent companies, there were in 1901 no tangible assets behind the common stock and perhaps two fifths of the preferred stock. The common stock represented hope. By a process of reinvestment of profits the Steel Corporation has been replacing hope with mines and plants. There is no doubt that the Steel Corporation has been a profitable venture; the Corporation's investment in tangible property on December 31, 1910, was greater by $504,928,653 than at the time of its organization in 1901, according to the figures compiled by the Bureau of Corporations. There may be doubt however as to whether the profits thus reinvested in the business had their origin in economies of operation. In many instances of course the company avoided the payment of profits to others by the operation of a coal mine or a railroad itself, but so far as it merely took profits which would otherwise have gone to others this does not represent economy from the standpoint of society. So far as the economies to be obtained from the integration of manufacturing processes were concerned, they were probably already obtained to the full degree by some of the larger and better equipped among the constituent companies. The chief economies possible were from the more economical subdivision of the business among the plants, by utilizing only the most efficient plants, and by dividing the work among the plants according to the nearness of the market. One writer in discussing this question states that "The conclusion would seem to be that the Corporation by virtue of its ownership of the cream of the ore and coking coal lands and of the iron ore railroads (not to mention its financial connections) had an advantage over its competitors but that this advantage did not demonstrate the superior economy of the trust form of organization."[1]

Other examples of vertical combination. Other examples of this type of combination are to be found in many lines of industry. We have already noted the integration of productive processes in the cotton industry, but in some instances we find integration of the production and selling functions. As a rule selling has been a specialized branch, separate selling houses handling the products of several mills, both in the domestic and in the foreign trade. Integration has, curiously enough, proceeded in two directions. In some cases the selling house has acquired mills, and in others the mills have either purchased controlling shares of stock in existing selling houses or have formed new ones. The fundamental cause of this, it has been asserted, is the growing competition offered by southern mills. Here again it is common to find consolidations of a joint horizontal and vertical type. The International Cotton Mills Corporation, which amalgamated several companies in 1910, controlled twenty-two mills, producing 3,000 varieties of cotton fabrics, and two selling houses.

---

[1] Eliot Jones, *The Trust Problem in the United States*, p. 219, *note*.

Conclusions. It seems unquestionable that both horizontal and vertical combinations offer the possibility of economies over and above those which can be obtained by the single large manufacturing plant, but the extent to which these have been realized in actual cases is exceedingly difficult to determine, for combination and monopoly have been so closely linked. It is however probably a fair statement of the case to say that where a combination has been conspicuously successful the profits have arisen in large part because of a monopolistic control of production which has enabled the combination to restrict the output and raise the price of the goods to the public, and not primarily because of superior methods of production or of marketing.

As a matter of fact the success of the large combinations of the United States has not been striking, when we measure success in terms of profits. Not only have profits been smaller than anticipated by the promoters of the consolidations, but in many instances they have been smaller than the aggregate profits of the constituent companies in the period before consolidation. Despite the economies which combination might have been expected to bring, the United States Steel Corporation during its first year of operation earned but 78 per cent of the amount earned by the member companies in the period immediately preceding the formation of the combination, and for the first ten years of its life its average earnings were somewhat less than its earnings during the first year. Professor Dewing has made a comparison of the earnings of a number of representative combinations with the earnings of the companies which were united and has found for the group as a whole that the earnings of the first year, as well as the average earnings for the first ten years of the existence of the combinations, were materially below the level of the previous earnings of the member companies.[1]

The weakness of combination. The failure of combinations to reflect economies of operation and management in their earnings seems to indicate that fundamental defects or weaknesses are by-products of size of organization. A combination may realize some of the economies of large scale management only to have these offset or more than offset by losses traceable to other factors. There are many losses to which a huge combination is peculiarly susceptible, but undoubtedly the chief weakness which organizations of huge size (including both large unit plants and combinations) have to face is the difficulty of finding men of the calibre required to direct the work successfully. It is possible to build up an organization which seems to work with effortless, machine-like precision, but examination will usually disclose a dynamic personality behind the organization, supplying the initiative, making the important decisions, and inspiring his colleagues with enthusiasm and confidence. The founder of almost any of our large industries has been such a man. His personal fortunes have been closely tied up with the fortunes of the business, and this has provided an effective stimulus. When the reins of management pass out of his hands into those of another, it is rare to find the same combination of ability and energy. If a son follows in his father's footsteps, even granted that he has equal ability, he will not have the same incentive, for not the least of the motives in the mind of the founder of a business is to leave a monument to himself. A hired manager can rarely have the same stimulus to activity, and, if his mind turns to venturesome improvements, his lack of financial control may make it impossible for him to carry out his projects. In Mr. Carnegie's time the scrap heap was in constant use. No matter how recently a machine had been installed, if there was one which was admittedly superior to it, Carnegie scrapped the other, and his competitors perforce followed his lead. It is difficult to believe that the scrap heap is as important an institution at present; there is a tendency manifest in many lines of business to block the trial of new machinery until the old has repaid its cost.

---

[1] Dewing, A. S., *The Financial Policy of Corporations*, vol. iv, Appendix.

Business ability of the highest order is required by all large scale production; it is especially essential for securing the economies of combination, and it is not sufficiently plentiful to give assurance that many large combinations will be able by more efficient production to undersell the large single units.

CHAPTER VIII

THE ORGANIZATION OF MARKETING

While our investigation thus far has been primarily devoted to the subject of production as a whole, we have nevertheless inclined to give a somewhat disproportionate share of our attention to that branch of production which is engaged in the extraction of materials and their fabrication, *i.e.*, in the creation of form utility, and our illustrative matter has largely been drawn from that branch of production. In so doing we have covered that part of the subject sufficiently for our present purpose. There remain to be considered the two other coördinate branches of production, concerned with the creation of time utility and place utility respectively, which are fairly covered under the titles of marketing and transportation. These topics will invite our attention in this and the following chapter.

Production to satisfy other people's wants. Production in the modern world under the régime of division of labor has become so organized that virtually the entire effort of every worker is devoted to the production of things for the use of other persons. It is only in the exceptional case that one produces anything for his own use. Such cases appear most frequently and regularly in the industry of farming. The ordinary general farmer secures a considerable part of his food from his own farm, and he sometimes obtains a few other things, such as firewood. But even he disposes of the bulk of his farm's product for the uses of other people, and with the specialized farmers the situation is even more obvious. A dairy farmer will take perhaps half a dozen quarts of milk for the use of his own household; whereas hundreds of quarts are taken to market. The truck farmer may use a few handfuls of his vegetables on his own table, while the truck loads go daily to market to satisfy other people's hunger.

On the great farms devoted to raising wheat, corn, and other cereal crops, on the cattle ranches, on the fruit farms, only an insignificant part of the product is devoted to the needs of the farmer or his farm laborers. Outside the farming industry the exceptions to the general rule are still more rare and unimportant. A shoemaker may occasionally make himself a pair of shoes; a lawyer may occasionally try his own case in court; the ordinary factory laborer may now and then use one or two out of the millions of units of the factory's product; now and then the salesgirl sells herself a pair of stockings or a hat, one out of the hundreds that pass through her hands. Many persons work all their lives making things for which they never have any use. Men are employed in making women's clothing; women, in making men's clothing. Books are printed and bound by workers who will never read one of them. Golf clubs and tennis rackets are made by men and women who do not play golf or tennis. It is the exception, rather than the rule, when a person makes use of the product of his own labor. And even when the worker does have as much use as anyone else for the product of his labor, his own need will be satisfied by an insignificant fraction of his total output. The worker in the sporting goods factory may be an enthusiastic golfer; yet he can use only half a dozen of the many thousands of clubs in whose manufacture he had a part.

Modern industry is highly specialized. The ordinary worker is a specialist and devotes all his energies to producing one thing or a very few things. But he wants hundreds of commodities and services, and for them he is dependent upon the efforts of thousands of other specialists working each in his own field. Thus we have the converse of the proposition with which this paragraph opened; namely, that virtually everything one has for the satisfaction of his wants must have been produced by other persons. Production for the wants of others and dependence on the production of others: these are the fundamental results of modern division of labor.

Bridging the gap between production and consumption. The statements of the preceding paragraph will, as soon as read, be recognized as nothing more than simple facts of common knowledge. And it is likewise common knowledge that, in spite of this radical separation between producers and users, the products of industry in some way get into the hands of those whose wants they were destined to satisfy. But the full significance of this situation is seldom realized, and the economic machinery by which it is accomplished is a complicated and marvellous thing. To recognize and understand the structure and functioning of this mechanism is one of the principal objects of the study of economics and is no mean task.

The function of trade. First of all it should be noted that the process of getting goods from makers to users involves trade; buying and selling. Goods are passed on from hand to hand, in countless streams of various sizes and directions, by means of trade. This relates not only to the finished product, but also to the various stages of production. The ranch owner sells cattle to the slaughterhouse or to the meat packer; the packer sells the hides from the slaughtered animals to the leather tanner. The tanner sells leather to the shoe manufacturer, and he sells shoes to the jobber, who sells them to the retail shoe dealer, who sells to the customer who is to wear the shoes. All of these exchanges were necessary to bring the one item of shoe leather from the ranch to the wearer of shoes. Just as many exchanges were required to obtain the cloth for the lining of the shoes, the thread for sewing the seams, the metal eyelets, the buttons or laces, etc. And if we stop to consider that the production of shoes required a cattle ranch, a leather tannery, a shoe factory, a warehouse, a store, railroads, motor trucks, typewriters, cash registers, banks, insurance companies, and so on indefinitely, the picture soon becomes too wide and complicated for us to grasp at a single glance. But it is evident that scores or even hundreds of transactions of purchase and sale are involved in the production of any everyday article of consumption.

The evolution of exchange. This was not always so. Exchange, thus insinuating itself into every stage of the processes of production and distribution, is a comparatively modern institution. Our understanding of it will be made easier by a glance at its historical development.

Among primitive peoples. Among primitive, savage, and barbarous peoples there is very little exchange. Many of their important possessions are owned in common by the whole tribe. So far as there is personal ownership, the things owned are generally the products of their owner's efforts in the hunt or in crude pastoral or agricultural pursuits. All are engaged in much the same occupations; there is little or no specialization and hence little reason for an exchange of products.

In civilized antiquity. The great nations of antiquity, such as Egypt, Persia, Assyria, Phoenicia, Greece, and Rome, rose to a high level of civilization. The exchange of goods in the modern sense however was scarcely present. Articles of everyday use, food, clothing, shelter, etc., were generally produced by the common effort of the tribe or clan or family. The products of such industry went into a common family or tribal store, out of which each received his share under the authority of the patriarch or chief. There was little individual ownership of wealth. The tribe or family was largely self-sufficient and also generally similar in its products and needs to the neighboring groups. There was, it is true, a considerable trade in foreign products, obtained through the agency of travelling merchants in exchange for the staple products of the nation. The goods thus received were generally in the nature of luxuries: silks and other delicate fabrics, wines, oils and perfumes, jewels, rare articles of food, slaves, etc. They were obtainable only by the few chiefs and leaders and persons of wealth and station. The ordinary economic life of the masses of the people was little affected by this trade.

In the Middle Ages. During the period of the Middle Ages the self-sufficiency of the small group, family, or tribe, was even more conspicuous than among the nations of civilized antiquity. The typical economic unit was the manor as it existed in Europe, an interesting type of economic organization which has been described in an earlier chapter (Chapter III). The reader will recall that there was practically no buying and selling among the various members of the manor and that the self-sufficiency and isolation of the manor were such that there was little trade with the outside. It will also be recalled from this previous chapter that the rise of the medieval town was the real starting point for the development of division of labor, specialization, and trade. The town craftsmen and merchants were specialists. There was division of labor between themselves and between the town and the neighboring country districts. Trade thus became a necessity. The residents of the town exchanged the products of their industry among themselves, and town and country exchanged with each other. Local products were even exchanged for the products of neighboring communities and for exotic merchandise from distant lands. Thus came the beginning of trade and with it the first examples of a marketing organization.

The traders. The ordinary retail trade of the medieval town was in the hands of the traders or merchants, whose status and organization have been described in Chapter III. They maintained small shops or booths, to which came the citizens of the town to purchase all kinds of articles, farm produce from the country district, products of the town's industries, and "foreign" wares brought in from other towns or other countries. It must be remembered however that the craftsmen also generally had small shops or booths in the front of their houses, where they sold the products of their crafts to the townspeople. They also made goods "to order" for their customers. Manufacture and trade were thus not entirely separate, especially at the beginning.

The merchant class did not confine its activities to local retail trade. Besides the local shopkeepers there were travelling merchants. They performed the function of carrying the surplus product of the town's manufacturing industries to other towns and other countries for sale. And they brought back to their home town a great variety of foreign products purchased in the neighboring towns or in more distant parts of the country or even in foreign lands. The travelling merchant of the Middle Ages has his parallel in the modern peddler, whose horse-drawn wagon, filled within and bristling on the outside with a bewildering collection of pots and pans, brooms and knives, dress goods and toys, and countless other articles, was once a familiar visitor in every American village and farmside and who even now may occasionally be encountered on the roads of the more backward country regions. He is a relic of the type of travelling merchant who in the Middle Ages furnished the one essential link between the products and the consumers of different towns and different nations.

The town market. The centre of the marketing organization of the Middle Ages was the town market, established in every town. Its principal function was to facilitate trade between the townspeople and the residents of the surrounding country. The market was a necessity because trade was still on so small a scale and so imperfectly organized. Today a farmer goes to town whenever he has produce to sell and finds it convenient to make the trip. He can always count on finding a wholesale merchant or broker to whom to sell his produce. He is also sure of finding retail stores always open and ready to sell him any ordinary thing that he may want in the way of local or imported products. In the medieval town there was not enough trading to call forth these facilities. It was therefore agreed between townspeople and country folk to set aside certain days upon which to meet at a certain place for the mutual exchange of their wares. Thus originated the "market," the "market-place," and the "market-day."

Markets very similar to those of the medieval town may be found today in many cities of Europe and the United States. For example certain streets or squares are often set aside for the produce market. Here gather each morning or on certain days the truck gardeners of the region surrounding the city, each with his wagon or motor truck loaded with vegetables and other produce. Here come the agents of the city grocery and provision stores, the commission brokers, the representatives of hotel and club stewards, and occasional thrifty housewives, to make their purchases. In some cities the market is inside a building, in which are rows of stalls, booths, and tables for the sale of various products. Certain specialized markets will often be found, such as fish markets, flower markets, etc. In some of the more backward countries of Europe and of South America the town market is of much greater importance than in the United States and bears a closer resemblance to the medieval institution. Often it is not continuous or daily, but is held only on certain days of the week, Sunday morning being a favorite time. As in the medieval towns, these markets serve as the meeting place for the exchange of town and country produce and are an important part of the economic life of the community. But, while this sort of market may be a fairly common and sometimes important institution in modern times, in the medieval town it was universal and all-important.

The medieval fair. Another aid to trade in the medieval period was the "fair." A fair was a sort of glorified market. The holding of a fair generally required permission from the king in the form of a special charter. Such charters were granted to feudal lords or high church dignitaries. Fairs were usually held only at important places, especially at cities which were junction points in the national routes of trade and travel and at the important ecclesiastical centres. They were generally held in connection with some gathering, frequently a religious festival, which brought large crowds of people together from a more or less widely extended territory. The fair might last for any period from a few days up to six weeks or more. Certain of the great fairs, held regularly at stated dates each year, came to be national or even international institutions, attracting great numbers of important merchants from far distant places in addition to the local traders and populace. The fair was an occasion for social diversion and merrymaking, as well as for business. Besides the booths and stalls of the traders, there were side shows, wild animals, trained dogs, magicians, musicians, and "freaks," resembling closely the side show attractions clustered about the main tent of a modern American circus.

The duke or bishop under whose auspices the fair was conducted generally endeavored to make the fair popular and to promote its growth in every possible way. He frequently secured for the visiting merchants exemption from the customary tolls and taxes, immunity from attachment of the person for debt, and a general freedom from the prevailing restrictions upon trade. Hence the fairs often came to be the only places where there was any approach to real freedom of trade, and they were consequently known as "free fairs." The noble patron also was active in compelling fair dealing among the traders, enforcing contracts, and protecting purchasers from fraud, by a great variety of rules and regulations. Special constables and even special sheriffs and special courts were present to enforce the rules of the fair. Fairs in which the freedom of trade was particularly liberal and the regulations especially fair and well enforced gained a deserved reputation and flourished accordingly. It should be recognized that the motives of the noble patron were not entirely disinterested. Special fees and taxes were collected from the merchants who traded at the fair. The patron of the fair was often granted a monopoly privilege, through which trading elsewhere than within the fair grounds was forbidden during the period of the fair. The revenue of a large fair was frequently a handsome sum.

These great fairs thus became the regular meeting places of merchants and traders from far and near. They furnished a clearing house for all the wares that entered into the national and international commerce of Europe, such as silks, fine fabrics of wool and linen, leather, skins and furs, jewels and various manufactured articles, wines, spices, foreign foodstuffs, animals, and slaves. Here the merchants from the neighboring towns could, in return for their home town products, purchase foreign wares to take back and sell in the local markets.

Of course the fair is by no means unknown to us today. In certain of the more backward countries, such as Russia, great fairs still exist, organized on the medieval model and performing a similar function. In the United States the fair, while still a trading device of some importance, has come to be more of an institution for exhibition and entertainment, though its advertising possibilities are fully appreciated by the modern manufacturer and merchant and in this way it serves the ends of trade. But after all the modern fair in most progressive countries is only a dim reminder of past greatness.

Transition to the present time. We have now before us the picture of the beginning of trade in Europe during the latter part of the period which we call the Middle Ages, say from about the year 1000 to the end of the fifteenth century, a period whose general economic organization we have studied in an earlier chapter.[1] How great an advance this was over the condition of family or tribal self-sufficiency and absence of trade which went before will be recognized. On the other hand it will equally be recognized that it is a long cry from the economic situation of the late Middle Ages to the infinitely complex organization of industry and exchange which, as we have seen, is the characteristic feature of the modern economic system. The record of how the régime of division of labor and exchange, painfully emerging from the lethargy of the Middle Ages, developed step by step into the modern system, is an interesting story. For our present purpose however it is not necessary to study the various steps of this evolution. We may pass at once to a survey of the marketing organization of the present day.

Present-day importance of marketing. Broadly speaking the subject of marketing embraces the study of the functions and activities of all those who are engaged in purveying or in assisting in purveying goods from those who extract and fabricate the materials to those who are to use the finished products, the consumers. We do not use the popular phrase " from producers to consumers," because we have had to recognize that the marketing agents are just as much producers as any others. We might, if desired, say " from the creators of form utility to the consumers." It is not always easy to draw a sharp line between the marketing agents and those who are creating form utilities, but that difficulty will cause us no serious embarrassment. On the other hand it is clear that our classification should include those who assist in financing, transporting, or insuring goods while in the process of transfer. In this chapter however we shall limit ourselves to the consideration of those who are engaged principally in buying and selling goods for the sake of the profit which may be obtained thereby; *i.e.*, to the traders.

---

[1] See Chapter III.

Marketing in the present day has attained to a degree of importance and attracted a measure of attention on the part of business men and others which are without parallel in any previous age. The fundamental causes of this development may be understood by reference to some of the features of modern industrial life with which the reader is now familiar. Present-day manufacturing is (1) capitalistic or roundabout, (2) on a large scale, and (3) highly specialized. Production is therefore for a market which is future, distant, and indeterminate in size. The village cobbler has no such problems as the modern large scale manufacturer in determining the scope of his activities or in disposing of his product. His neighbors drop in and order shoes, and he quotes a price which he knows will be remunerative. The manufacturer of shoes in Lynn is not making them for Lynn only, nor yet for the people of Massachusetts or even of the United States. His market is limited by no geographical bounds; wherever shoes are worn he either does or can sell them.

For months previous to placing shoes for spring and summer on the market he must have been engaged in manufacturing them. But in order to start manufacturing, machinery and equipment must be in existence and must therefore have been in the process of production for several years. With such an interval between the initial stages of production and the satisfaction of the wants of the consumer there is inevitable opportunity for miscalculation. We have observed in a previous chapter that the factor which determines how much the shoe manufacturer will produce is the relation between the price he can get for the article and the prices which he has to pay to get the shoes made — the prices of leather, labor, etc. But prices change, and after he has started his production for a given season they may drop to an unprofitable level. Any information which will shed light on the probable condition of the market will have its influence on prices and will therefore enable price to be a more effective regulator of production. And if this is true of a good which is capable of yielding immediate enjoyment, how much more so of a good which is intermediate in the production of that good. Knowledge of markets and marketing conditions is therefore of vital importance in bringing about a nice adjustment of production to consumption.

Marketing methods have assumed a greater degree of importance relative to the other branches of production. In the latter most of the processes and methods are more or less the common property of all producers, although each new improvement may give a temporary advantage to one producer and is therefore eagerly sought. Yet there is a greater degree of stability in manufacturing than even thirty years ago. Producers today are less jealous in guarding their business secrets than formerly, and independent institutions of research, such as the Harvard Committee on Economic Research and others, make public the methods of internal organization which have proved most successful. On the other hand the producer is coming to rely more and more on securing a wider market for his goods by effecting economies in his method of distributing his product, and he is further impelled to do so by the fact that the only way that may be open to him to reduce the unit cost of manufacturing is to expand present markets or to develop new ones. If he can market his goods more efficiently he can reduce the price; if he reduces the price he can sell more goods; if he can sell more goods he can secure economies from large scale production — and so the cycle starts over again. Hence we find the manufacturer ever ready to try out innovations in marketing methods, both in domestic and in foreign trade.

The consumer is no whit less interested in marketing problems. The difference between the price which the farmer receives for his milk or the manufacturer for his automobile tires and the price which the consumer pays the retailer sometimes seems to him outrageous. He views the toll which the middlemen exact as extortion and wants something done about it. His interest, be it noted, is of less importance in shaping marketing organization, but that it is not negligible can be seen in the phenomenal development of the mail-order houses.

Marketing functions: assembling. We have defined marketing in general terms; it will be useful to classify and describe the functions which can properly be called marketing functions. The first of these is *assembling.* This involves the collecting of goods from a variety of sources. It may be a simple matter, as when the bulk of the production is handled by a few large manufacturers, or it may be more difficult and call for more painstaking work, as when the good has to be collected from a large number of small scale producers. Tea in India and Ceylon is produced on large plantations owned by large companies; in China it is grown by small tea planters, and the task of collecting it in the latter case is infinitely more difficult than in the former. To perform this function successfully, and by this we mean profitably, a careful study of markets and marketing conditions is requisite in order to determine what factors may influence the volume of production and therefore the price. Climatic changes such as prolonged wet weather or prolonged drought in any large producing area determine the keenness of competition between those performing this function and influence the price which they are willing to pay. In 1891 there was a total failure of the wheat crop in South Russia, and in France a serious harvest shortage, while the United States enjoyed the largest crop in its history up to that time. The effect of the disclosure of this situation on the activities of those buying wheat from the farmers can easily be imagined. Local disturbances, which may have a decided influence on production and trade, have to be considered. Thus the revolution in São Paulo, Brazil, in 1924 affected the price of coffee. Labor troubles may tie up the productive or transportation system in an important producing country and cause a temporary realignment of the important producing countries. The strikes which have been of frequent occurrence in our coal mines have diverted coal from other producing countries to this country. The one who most successfully performs this function is the one who foresees the event before the crisis occurs and who, in the event that habitual sources of supply are closed to him, secures a stock from other sources while yet there is time.

Storing. Goods cannot be passed on directly from manufacturer to consumer except in a few instances; in the majority of cases someone else must store them while they are being gradually used up. Few consumers for example are willing to buy a winter's supply of flour. They prefer to buy it as they need it, and, as was seen in the hoarding of sugar during the World War, only fear of privation or of extraordinary rise in prices will induce them to do otherwise.

The assumption of risk. It is a rather obvious statement that all business involves risk, but its truth is frequently overlooked, and the logical deductions to be drawn from it are little understood. In marketing there is the risk of loss by fire or theft, the danger that it may be impossible to sell goods at a profit, the possibility of loss through uncollectible accounts. Someone must bear these risks. The early merchant assumed most of them; the present middlemen try to shift them to others when they can. Insurance companies are organized to assume the risks of fire and theft and some companies will, at a price, shoulder any risk. Others remain on the middlemen or on other producers.

Rearrangement. The wool which is collected from a number of small ranches must be sorted, graded, and packed for shipment. Before it is used it may be combined with other lots and perhaps be regraded and repacked. Wheat received from the farmer must be cleaned and graded; coffee shipped in bulk from Brazil undergoes the same treatment; and even manufactured goods are not always ready to be sold to the consumer in the form in which they are received from the manufacturer, but must be marked and packed. This task of grading, sorting, and packing is called rearrangement and is by no means an unimportant part of the middleman's functions.

Selling. The consumer must be told of the existence of a product, and in many cases he must be educated into wanting it. Furthermore, he must be induced to believe that goods put out by Jones & Company are much superior to those put out by Smith & Company. Personal solicitation, solicitation through the mails, and advertising all play an important and, it may be added, an expensive rôle.

Transportation. This is a function of very great importance, for the degree to which transportation facilities are developed determines the size of the market and to some extent the nature of the marketing organization. In this section we shall assume that adequate transportation facilities are provided by other agencies than the strictly merchandising middlemen. Only where such middlemen carry on this function themselves, as in trucking, shall we here treat of this function.

Financing. The system of deferred payments is almost invariably found in one form or another in marketing goods. Someone in the chain of producers will be found unable to make immediate payment for the goods he purchases. The wholesale merchant may have sufficient resources of his own to extend credit to the retailer, but frequently he can do so only if another comes to his assistance.

The extension of credit is still an important function of the middleman, but in our modern organization it has in large measure been transferred to specialized agents. The commercial banks, the discount houses, the note brokers, and others are able by gathering together the savings of many to extend credit to the merchant or the manufacturer and thus facilitate the continuous flow of goods to the consumer. The part which they play in our organization is one of great significance, but the consideration of the nature of their operations must be deferred to later chapters in our study.

Marketing agents. In a very brief way we have examined the marketing functions; we shall now glance at the agents engaged in distributing the products of industry in order to see which ones perform these different services and how each class contributes toward helping the consumer get what he wants when he wants it.

The retailer. The retailers form the largest and in some respects the most important group of middlemen. They are in closest touch with the consumers and should know the probable quantities of different goods which may be sold at various prices to their customers. In any case the other groups of middlemen have to rely on them to gauge the market, and they look upon the orders which the retailers pass on as one of the most reliable indices of business conditions.

The old-fashioned retail store, which still exists, is compelled to deal in almost every known article from groceries to millinery. In a small village one is forced to get a large enough volume of business to yield a profit by appealing to many different needs. With the growth in the volume of trade it has been possible to concentrate on one or more lines of business, and the modern specialty shop has appeared. Thus the selling of clothing and food in the same shop has pretty generally been given up, and even in the selling of clothing there has developed specialization in men's and women's clothing, to mention only two examples. There is a real advantage to the retailer and to the public in such specialization if the volume of business is large enough to warrant it, for the retailer has the chance to buy larger quantities at a lower unit price, and the consumer gains, possibly through lower prices, but certainly through more efficient service and a wider range of goods from which to choose.

The modern department store represents a sort of glorified country store in the bewildering array of goods which it carries, but it has gone beyond the scope of the ordinary retailer and has assumed some of the functions of the wholesaler in that it may deal directly with manufacturers and may regularly send its own buying agents abroad. It may even engage in the manufacture of some of the goods which it sells. Besides the department stores we have other examples of the modern development in retailing: the chain stores, such as The Great Atlantic & Pacific Tea Company and the United Cigar Stores, which are simply small retail stores under a centralized management, and the mail-order houses, as Sears, Roebuck & Co., which attempt to reach the retail trade through catalogues and mail campaigns. These large units can probably buy more cheaply in large quantities than the small store, and they can use their labor force more efficiently.

In spite of the economies of the forms of retail organization mentioned above, the small store persists and seems likely to remain an important part of our organization. Acknowledged by most investigators to be the least efficient of all the marketing organizations, charging in many cases a higher price than elsewhere, it persists because many people prefer to shop there. The only explanation there can be is that it gives a service for which society is willing to pay and appeals to a feeling which is widespread. It is the feeling to which appeal is made in the "Neighborhood Store" signs placed so conspicuously in street cars in many cities. The ordinary individual likes to trade where he is known and where the proprietor without being told gives him the particular brand of cigar he uses or the morning paper he habitually reads. It soothes his vanity. He furthermore expects and gets service which is not given to every newcomer. The proprietor of a drug store does not feel quite so annoyed when one who wakes him in the early hours of the morning for medicine turns out to be an old customer; the owner of the garage where one has his repairs attended to drops work on the car of the stranger and attends to the wants of his regular customer.

The retail store of moderate size usually performs few or none of the tasks of assembling or rearrangement, but buys the goods in the form desired. To some extent goods must be kept in stock, but the careful retailer keeps his stocks as small as possible in order to avoid the risk of price fluctuations, to be sure that his stock of goods is fresh, and to keep his investment in goods at as low a figure as possible. He may do some financing through the extension of credit to customers, although success in retail trading is based to a great extent on cash transactions. The margin of profit is so small that the retailer cannot generally run the risk of bad debts. On the other hand he is frequently in debt to the middlemen from whom he has bought goods, and he may stay in debt from one year's end to the other, except possibly for a few days when a complete settling of the account may be required. If he delivers goods to the customers' homes, transportation should be included in the list of his activities.

The department store and the other large scale retail agencies perform more of the marketing functions than the ordinary retailer. Assembling and rearrangement may be fairly important features of their business, selling is usually more highly developed, advertising and mail campaigns play a rôle of some importance, and the sales force of a department store is given training in the arrangement of goods in the store, in the line of goods being sold, and in sales methods generally.

That the retail stores offer services which the public wants and for which it is willing to pay is so patent as to need little elaboration. They seem almost indispensable as a means of spreading goods before the public eye. Even in the relatively rare cases where the manufacturer attempts to reach the consumer directly, there is not an elimination of the retail store element — it is merely assumed by the manufacturer in addition to his other operations. The Douglas Shoe Company for example runs a number of retail shoe stores for the exclusive sale of shoes of its own manufacture. But even in the event that no formal retail stores are operated, there must be an elaborate retail department in connection with the factory and considerable expense incurred in soliciting patronage and shipping the goods ordered.

The wholesale merchant. There are many wholesale agencies of one sort or another, such as the wholesale merchant, the commission merchant, the manufacturer's agent, and the broker. The wholesale merchant purchases goods outright from the manufacturer and sells them to the retailer, relying on the difference between the buying and selling prices to make his profit. Such a wholesaler is sometimes called a *jobber*. A jobber formerly was a man who bought in odd lots and sold to occasional customers, whereas the wholesale merchant dealt in large quantities and with an established clientele. At present there is no valid distinction to be drawn between these two terms.

The wholesale merchant performs many of the marketing functions which the retailer refuses to assume. He may collect goods from widely scattered sources, he usually has to rearrange them, and he may keep large quantities of goods in stock. Some wholesale houses on the other hand carry a very small stock of goods, forcing the manufacturer to assume this burden. This means that not only the cost of carrying the stock, including warehouse charges and the interest on the money tied up in the goods, but also the risk of price fluctuations is borne by the manufacturer. Where this situation exists the wholesale merchant usually refers orders to the manufacturer with the request that he make shipment from the factory. The manufacturer tries as a rule to keep his goods moving from the factory to the warehouses of the merchants as rapidly as possible; he prefers to manufacture on the basis of orders given him and not to run the risk which is inherent in manufacturing for an indefinite market.

Selling forms an important part of the wholesale merchants' work. Although some houses play a passive rôle and expect their customers to come to them, competition has forced most houses to adopt aggressive selling methods, and we find them employing a large number of high grade salesmen and advertising extensively in the trade papers. Some wholesale houses buy unbranded goods from manufacturers and put them out under their own trade names or brands; others handle the competing branded goods of manufacturers. In the former case the wholesaler has an incentive to push the sale of his own goods at the expense of competing brands, while in the latter the only preference which he will show will be determined by the relative profit in handling one brand as compared with another.

In most cases the wholesaler pays the manufacturer cash or its equivalent for the goods and extends credit to the retailer, thus performing a financial service of considerable importance. It is becoming more common for the wholesaler to assume transportation functions also. Many wholesalers ship the goods from their warehouses in their own trucks, using the railways only for distances which cannot be covered in a day's run.

The list of the functions which the wholesale merchant thus performs should convince us that he renders services which neither the retailer nor the manufacturer could handle so successfully. As a specialized middleman — and he may specialize not only in handling goods in large quantities but also in one commodity or in a group of closely allied commodities — he has a knowledge of markets and of men which the retailer and the manufacturer could not acquire except at greater cost. The task of obtaining all the goods which the retailer requires from numerous manufacturers, widely separated geographically, would be a difficult and expensive one for him. It is much simpler to order all that he requires in a given line from the wholesaler's salesman, effecting a saving in time and in clerical expense. The wholesaler receives goods in carload lots from the manufacturer, breaks them up into convenient size, and sends them by truck to his many retail customers. This means a great saving in freight charges over what would be necessary were the manufacturer to ship in small lots directly to the retailer. On the other hand the wholesaler is in a better position than the manufacturer, who makes only a few products, to estimate the probable demand for goods, for he has a broad view of the demand for all articles in his line and is less likely to be led astray by a change in the demand for a particular article. In summary we may say that the wholesale merchant relieves the retailer and the manufacturer of many of the costs and risks of marketing, thus permitting each to specialize in that branch in which he is presumably most efficient, and that he performs his services at a lower cost than either the manufacturer or the retailer could.

Other wholesale agencies. Goods may proceed directly from the manufacturer to the wholesaler, to the retailer, and to the consumer, but manufacturers in many cases have not been wholly satisfied with complete dependence on the wholesale merchant. The proper and economical marketing of their goods has seemed so vital a problem that other agencies have been used at different times to supplement the work of the wholesale merchant or to supplant it. In some lines the manufacturer has tried to eliminate the wholesaler entirely and to deal directly with the retailer. This may be feasible if the manufacturer makes a large enough number of products so that it will be worth his while to employ an expensive sales force to cultivate the retail trade, but a manufacturer making a single product can scarcely find such an organization profitable unless he sells exclusively to stores operating on a large scale or unless there is a profit of considerable magnitude to be derived from the sale of each article. In most cases where the manufacturer is dissatisfied with his wholesaler the best thing for him to do is to seek another distributing agent or to stimulate the demand for the product through advertising.

The *commission merchant* is sometimes used to handle the entire output of a producer or a group of producers. He receives goods on consignment; that is, they are sent to him to be sold, and he is paid by a commission of a fixed percentage of the amount of the sales. The commission house is of considerable importance in the textile trade, frequently playing an important part in the financing of the mills. The *manufacturer's agent* differs little in theory from the commission merchant. He does not purchase the goods himself, and he is paid on a commission basis. He is however usually restricted in territory and sells the goods under the producer's brand, whereas the commission merchant takes the entire output, sells it where he can, and usually marks the goods with his own brand.

The *broker* is another agent who may appear if the volume of sales is sufficiently large to make his services of value. A broker is an agent representing either the purchaser or the seller and receiving a commission for sales effected. He differs from a commission merchant, who also stands in the relation of agent to principal, in that he does not have actual possession of the goods and does not conclude transactions in his own name but in the name of his principal. He rarely represents any particular manufacturer, but places orders given him with the manufacturer who will quote him the best price.

Markets. Thus far, both in our investigation of the early markets and in our analysis of the modern marketing organization, we have not found it necessary to formulate a precise definition of a market. Briefly a market is a place where buyers and sellers effect exchanges of goods. There may be but one market for a particular good or a number of unconnected markets, but in the case of goods of importance in our industrial life there are likely to be several markets intimately connected with one another. In the trade in raw cotton each local selling point is a market, but there are others occupying a more prominent position, which we call *primary* markets, and still others standing between local and primary markets, which we call *secondary* markets. Thus Liverpool is the best example of a primary market in the cotton trade and

Chicago in the wheat trade. But any place where the forces of demand and supply are so focused as to bring about the determination of a price is a market.

The produce exchange. The produce exchanges are simply markets organized on a more elaborate scale than most markets, having more highly refined facilities for trading and embracing a larger number of buyers and sellers. The produce exchanges started as voluntary associations of men dealing in one or more commodities and have continued on that basis. The exchange itself does not deal in the commodities, but provides the facilities, formulates the rules governing the sale of commodities, settles trade disputes, establishes standards for grading the commodities dealt in, gathers and distributes statistical information about crops and markets, and through the ticker service gives immediate publicity to sales and prices on the market. Only members of the exchange may deal on the floor of the exchange. These are brokers in most cases; that is, they are buying and selling for others, but there are dealers as well who buy and sell on their own accounts.

A feature of the exchanges, marking a great difference between them and the city vegetable markets, is that the wheat or cotton is not brought into the exchange but is kept in warehouses which may be miles away from the exchange. The sale is made on the basis of samples which are displayed or of definite grades which are standard and known to the trade.

The exchanges also offer a market for future goods. Contracts for the sale or purchase of wheat or cotton yet ungrown are freely made. This serves to make it possible for a manufacturer using raw cotton to quote prices on future deliveries of cotton yarn or cloth and relieves him of some of the risk of manufacturing. It also facilitates speculative dealings, a feature of considerable importance and value in our marketing organization which will be treated at length in a later chapter.

Not every good can be sold in an organized market of this type. Perishable goods such as fruits and vegetables are obviously ruled out. A good which does not lose value over a short period of time through the natural process of decay or deterioration, which cannot be rapidly reproduced so as to upset all calculations as to the total quantity in existence, and which can be easily graded and sampled is the most suitable. A further requirement is that it be of sufficient importance to occupy the attention of a specialized body of men. Such a commodity as wheat or cotton can be dealt with most effectively on the exchanges.

The services offered by the produce exchanges are very real. All the forces of demand and supply the world over are centred in the exchange through almost instantaneous communication with other exchanges. A change in the prospects of the Russian wheat crop is instantly reflected in the price of wheat in the Chicago Board of Trade and, in the manner already described, influences the consumption of wheat and the prospective planting of wheat in all the producing countries. As the market area becomes more extended and as speed and ease in exchanging goods are promoted, there is a smaller range of price fluctuations within a given period. Price is understood to reflect the conditions of consumption and of production and to regulate them both. Organized exchanges permit the determination of prices which more accurately reflect the conditions of consumption and production than in an unorganized market and therefore promote the more effective regulation of production and consumption.

Services of the middlemen. There are other institutions and methods which exist to expedite the flow of goods from the producer to the consumer. The selling organization of almost every important commodity includes some feature peculiar to itself. It is not our purpose however to attempt a detailed survey of marketing systems, but simply to obtain a general view of the organization from which to form some opinion of the place which marketing holds in our economic life and of the services which the middlemen perform.

It is probably clear that the middlemen do perform functions which society considers essential to its well-being and for which it is willing to pay. It will further be readily admitted that middlemen can perform these functions more economically and more satisfactorily than either the other types of producers or the consumers. But the questions will doubtless be raised: why so many middlemen, and why such special middlemen as cotton brokers or manufacturers' agents? Are they really necessary or are they social parasites?

In particular cases there can be no doubt but that the organization could be somewhat simplified and goods sent from the producer to the consumer at a lower cost than at present. In general however, considering the marketing organization as a whole, the conclusion is irresistibly forced on one that the complexity of the market organization of the present is the result of various attempts to market goods cheaply and expeditiously and that it represents the most economical solution of this problem which has yet been found.

Evolution of marketing. It is essential for the reader to have clearly in mind the realization that, under our economic system, the relationships in production and distribution as well as in all other phases of economic life are not stable. They are subject to constant change, in response to changes in the world of which they are a part. The organization which is competent to handle trade in one industrial stage is unequal to the task in the next, and modifications set in, slowly to be sure but nevertheless certainly. The constant strivings of the consumer to get goods as cheaply as possible, of the producer to sell his goods at prices which will be most profitable, and even of the middleman to increase his profits have the effect of a great natural force impelling changes in the organization which will eliminate unnecessary complexities. There is of course scope for governmental regulation in some matters — competition does not always bring the results which are generally desired — but this can be limited to defining the rules for playing the game; in other words, determining the plane of competition.

Our present organization is the result of changes in production and consumption consequent upon the Industrial Revolution. The old organization, which was considerably simpler than the present, was incapable of handling the vast quantities of goods which could be produced in factories equipped with power-driven machinery. Changes in the marketing organization were therefore forced upon the merchants. Large scale methods in production brought about large scale methods in marketing. The development was largely in the direction of greater specialization on the part of the marketing agencies and the creation of new agencies to help handle the increased volume of goods.

An example of this development will help us to understand it. In the year 1785 the cotton industry, as we have already seen, was in its infancy. The importation of raw cotton into Liverpool was a trade of no great significance. Cotton was imported by general commission merchants and sold by them to small dealers, who resold it to the manufacturers a bale or two at a time. As time went on the importation of raw cotton, particularly from the United States, increased prodigiously. It became so important a branch of trade that a highly specialized organization developed to handle it, and by 1840 we find the situation entirely changed. The actual importation was handled in the main by merchants who bought on their own account, in many cases maintaining buying offices in the United States. The general importer gave place to the specialized importer. The dealer had been forced out of existence and his place was taken by two sets of brokers, one representing the manufacturer and the other the importer. In 1785 there were brokers who sometimes effected sales of cotton, but they were general brokers, whereas these new brokers handled nothing but cotton. At the same time special facilities for trading in cotton had grown up, banks which almost exclusively financed purchases of cotton, transportation and warehousing facilities devoted to cotton alone, etc.

This highly specialized organization in the cotton trade — and there are other examples — grew up in response to the demand that the large volume of cotton be handled more effectively than in the past, and that it was more cheaply handled is evidenced by a reduction in the brokerage charge from one per cent to one half of one per cent. Inefficient agents in the organization, such as the dealers, were swept away, much as they may have desired to remain and get a profit. The importing merchants, the brokers, and even the bankers could operate more efficiently when they had but one commodity to handle than as general importers or brokers. They acquired special knowledge of the producing areas, of the foreign markets, and of the local needs of manufacturers. They became experts in their knowledge of the raw cotton itself. This facilitated grading cotton and selling it by sample and saved much time for both seller and buyer.

Specialization, with its economies in selling, cannot come about unless the volume of business in one commodity is large enough to occupy all of a man's time. The medieval merchant assumed all of the functions of the middleman. He bought goods from producers at home or abroad and retailed them himself. As the volume of trade increased some of these merchants sold goods to others who also sold them at retail, but in the course of time it occurred to someone that he had enough work to keep him busy if he confined himself to gathering goods from the manufacturer and selling them to the retailers. Thus the wholesale merchant, specializing in buying large quantities of goods, came into being. As we follow the course of history we see specialization developing with the growth in the volume of trade. There is a division of the functions of the middlemen among an increasingly large number, and some of the functions, such as selling, are performed by a number of agents, each one attending to only a fraction of the task.

This is of course simply division of labor in marketing and an illustration of the fact that the extent of division of labor is limited by the size of the market. We have gone into it in some detail because its significance in this field does not seem to be so patent as in the field of manufacturing. Yet it is essentially the same thing, and the economies which result from it are of the same character.

CHAPTER IX

THE ORGANIZATION OF TRANSPORTATION

Frequent reference has been made in the preceding chapters to the place of transportation in our economic organization. We have seen that transportation is primarily a function of the middleman but that it has grown to be of such importance that specialized agencies have developed to handle it. We have also seen that the development of transportation facilities is fundamental to the extension of division of labor in the sense that specialization in the product for which a region is best fitted (geographical division of labor) is impossible unless there is a profitable market for the goods produced.

We shall be chiefly concerned in this chapter with the transportation problems of the United States. No other country offers so excellent a field for the study of the influence of transportation on economic development, either in the matter of the localization of industries geographically or in the development of great centres of trade and population.

Conditions before the nineteenth century. Let us remember that at the close of the Revolution the population was confined chiefly to the eastern seaboard. Settlements began at the mouth of a river and extended inland as far as the river was navigable but proceeded only slowly beyond that point. Access to the fertile plains of the Middle West was impeded by the great Appalachian system extending from Vermont to Alabama, covered with heavy forests in many places and offering relatively few natural passes or gaps. Of these, some which offered the least natural difficulties, such as the Mohawk Valley, were shunned for many years because of the Indian menace.

Communication between states was rendered difficult because of the lack of roads and the fact that few of the rivers were navigable for many miles. Coastwise sailing vessels handled such commerce between states as there was, but this method of transportation was both hazardous and costly. There was, it is true, some travelling by road, but it was dangerous and time-consuming, and the transportation of goods by land was prohibitively expensive.

The lack of means of internal transportation was not felt so keenly before the Revolution as after it. The colonies were all much more dependent on foreign than on domestic trade, and we find it quite natural that attention should have been turned to the building of a merchant marine. Not only was there a considerable volume of goods to be exchanged with foreign countries, but local conditions were such as to foster the growth of ocean transportation facilities. Innumerable good harbors, a plentiful supply of raw materials for shipbuilding at the water's edge, and, in New England at least, the need for some more profitable outlet for the energies of the people than that afforded by the rocky and sterile soil led to an interest in the sea and a great development of the carrying trade.

In 1790 not more than five per cent of the American population was settled in the regions west of the Appalachians, but the tide was moving in that direction, and in the first years of the nineteenth century thousands crossed the mountain ranges to take up their abode in the new regions. It is natural to find that this situation led to dissatisfaction with the means of internal communication. The settlers needed a market for their produce and needed the goods which the eastern states could provide either through their own manufactures or through import from abroad.

The era of road-building. Attention was first directed to the building of a network of roads connecting the various parts of the country. Before the Revolution local roads between some of the towns had been built in all of the states, and some through roads were in operation, but the widespread interest in road-making came later.

Turnpike companies were chartered and were allowed to charge tolls of passengers and freight to compensate for the expense of building and maintaining the roads. Some of the states gave liberal aid to the companies. Pennsylvania had subscribed about $2,000,000 to fifty-six turnpike companies by 1822, about a fifth of this amount for the building of bridges, and by the end of the first quarter of the century had about 2,000 miles of improved turnpikes.

The interest of the federal government in the building of roads is seen in the construction of the Cumberland Road. There were three main roads leading across the mountain chain, a road up the Mohawk Valley into western New York, the Forbes or Pennsylvania Road from Philadelphia across the Allegheny mountains to Pittsburgh, and the Wilderness Road up the Shenandoah Valley through the Cumberland Gap to the Ohio. There was however need for another and better road to the Ohio Valley, and in 1806 Congress authorized the building of a wide, well-surfaced, and well-drained road. Starting at Cumberland in Maryland it was eventually pushed on to Vandalia, Illinois, and it was anticipated that it would go even farther, but the development of the railroad made further extension seem unwise. The Cumberland Road was opened to traffic at Wheeling on the Ohio in 1816, and it served to connect the Ohio and Mississippi valleys with the East.

The building of turnpikes, however desirable as a means of linking up the navigable streams and promoting internal communication between the several states, did not solve the problem of the transportation of the wares of commerce. It was true that a wagon could now pass where before only a horseman could with difficulty pick his way and that the roads could be used in most seasons. But the roads were by no means perfect, and many of them were impassable at some periods of the year. The greatest difficulty however was the cost of transportation. The cost varied of course according to the character of the road and of the country through which it passed, but ten dollars per ton per hundred miles is generally admitted to have been about the average charge. Bulky goods could not stand this charge. The cost of transporting wheat from the Ohio Valley to the seaboard would have been far greater than the value of the wheat itself.

The navigable rivers. So far as internal commerce was concerned the navigable rivers were more important than the turnpikes. We have noted that most of the rivers of the eastern seaboard were navigable for short distances only, but there were some important exceptions. The Hudson River served as an artery of commerce for the region north of Albany and partially determined the importance of New York as a commercial centre. The Susquehanna, running through the western part of New York and Pennsylvania, directed the flow of goods to Baltimore. Even the Connecticut River, which today is considered to be navigable only as far as Hartford, was then a trade route of considerable importance and was open to craft of small draught for two hundred miles or more.

The most important system of waterways however was found in the Middle West in the Mississippi River and the other rivers flowing into it. For a period of about ten years (1784–1795) navigation on this river was blocked by the Spanish at New Orleans and there were no markets open for the produce of the newly settled regions. The opening of the mouth of the Mississippi in 1795 gave the first market, and even before 1800 a considerable river traffic had already sprung up. Flatboats loaded with local produce of every description were sent to New Orleans, where the boat and the cargo were sold. In the early period the crews took ship for Philadelphia or Baltimore, where they purchased the manufactured goods needed at home and transported them over the mountains by wagon. Later it was customary to make the return journey by land following the course of the river on foot, but until the advent of steam few attempted to come up by boat against the strong currents.

The invention of the steamboat and its introduction on the western rivers meant an immediate and immense increase in the volume of commerce carried, especially since the question of the right to the use of the lower Mississippi had definitely been settled by the Louisiana Purchase in 1803. The steamboat was first used on the Ohio River in 1811, and in 1817 the first voyage from New Orleans to Louisville was made. The following years saw a rapid increase in the number and in the capacity of the steamboats plying on the rivers, until by 1847, it is stated, the steam tonnage of the Mississippi Valley exceeded that of the whole British Empire.

The effect of this change was an immediate increase in the volume of traffic. In 1816 the receipts at New Orleans amounted to less than ten million dollars; in 1818–19, after the steamboat had proved its practicability, they amounted to nearly seventeen million dollars, and most of the increase can be attributed to produce brought down the river. There was also a great saving in the time required for shipment. It took a flatboat about three months to go from Louisville to New Orleans, while in 1817 the trip was made by steamboat in twenty-five days, and ten years later in eight or nine days. Similarly a reduction in the cost of shipments followed in the wake of the improvements in river transportation facilities.

The steamboat did not put an end to shipment by flatboat, for in the beginning it was not any cheaper. In fact it made shipment by flatboat simpler, as the crew could return by steamboat instead of taking the long and wearisome journey by land.

Cities located at strategic points on the rivers became important shipping points for the produce of the surrounding country. Pittsburgh, Cincinnati, Louisville, and others grew in population and importance. Cincinnati became the centre of the pork packing business. Formerly it had been customary to drive thousands of hogs over the mountains to Philadelphia or Baltimore each year, as the hogs themselves furnished the only known means of transportation, but the opening of the river made it possible to ship bacon and salt pork to New Orleans for sale there for local or foreign use. Louisville was the centre of a thriving tobacco trade, and Pittsburgh served as a gateway between the East and the Ohio Valley and was an entrepôt for the wares of the East.

The canal era. While the improvement of the roads and the excellence of the western water routes stimulated commerce, there was need felt for a cheaper means of transportation between the waterways, particularly between the eastern and the western parts of the country, than was provided by the turnpikes. The solution to this problem was found in the building of a system of artificial waterways, linking up various navigable streams and providing a continuous means of water transit from inland to tidewater. The construction of these canals was undertaken partly by private companies, which expected to get a profit from the collection of tolls, and partly by the state governments.

Some canals were projected and built before 1800, but the era of canal building does not really commence until after the War of 1812 and ends about 1850, when the railroad was definitely proving its superiority. The first canal of commercial importance was the Erie Canal, built by the state of New York, connecting the Hudson River with Lake Erie. The nature of the Mohawk Valley, through which it runs, is such that it offered the least difficulties in construction of any possible route across the mountain ridge. The canal was begun in 1817 and finished in 1825, when a boat first made the trip from Lake Erie to New York City.

The Erie Canal was a success from the very start, and in ten years the tolls collected had paid for the entire cost. The freight rates from Buffalo to New York fell from $100 per ton to less than a quarter of the amount, and the time consumed in making the trip decreased from twenty to eight days. Wherever a natural waterway connected with the canal a thriving city sprang up. New York itself acquired its position of commercial supremacy very largely through the influence of the canal in diverting traffic to it. It is interesting to note that at one time it was thought that the region around Chesapeake Bay would be the centre of population because of the shipping facilities offered by the indented and sheltered coast line and the large area which could be served by the Susquehanna River. Western New York was in fact more closely linked with Baltimore in this period than with New York City, as the Susquehanna River afforded the cheapest means of shipping produce to the market.

Other canals were built connecting the Erie with Lake Champlain and with Lake Ontario, and still later (1855) the construction of the Sault Ste. Marie Canal connected Lake Superior with the other lakes. In 1832 the Ohio River was connected with Lake Erie and shipment of goods from the Ohio Valley by the way of the Erie Canal instead of by the Mississippi was made possible. Thus New York could serve as the terminal point for shipments from the fertile western part of the state and from the rich regions around the Great Lakes. A large volume of goods was still sent down the Mississippi River, larger indeed than the volume going over the Erie Canal and the Hudson River. The latter route carried most of the produce from the northern part of Ohio and the regions bordering on the lakes and was especially important in providing a means for shipping goods into the interior, far surpassing the Mississippi in this respect.

The success of the Erie Canal and its feeders stimulated the building of canals by other states. Some were local in character, but others were intended to meet the growing competition of New York as a commercial centre. Boston and Philadelphia were especially concerned with the threat to their positions in internal and external commerce. The former proposed to build a canal across the state to divert traffic to its port but never constructed it. A canal was however built from Philadelphia to Pittsburgh on the Ohio. A number of portages by horse railway were necessary to cross the mountain ridge, but it was completed in 1834 and remained in successful operation for many years.

We cannot attempt to describe all of the projects of canalization which are to be found in this period. Many of those built for local purposes, such as the Middlesex Canal connecting Lowell and Boston, were profitable and continued in use until forced out of existence by the competition of the railroads. They served an exceedingly useful purpose in their time in promoting the exchange of goods between various places. The importance which contemporary opinion attached to canals is reflected in the assistance which many of the state governments gave to canal building. Up to 1838 state debts which had been contracted to assist canals amounted to over $60,000,000. Pennsylvania led with a debt of $16,000,000 for canals, and New York held second place with a debt of about $13,000,000.

Development of the waterways and division of labor. Before proceeding to the next period in the development of our transportation facilities — the railway era — it will be worth while to consider the relation of the geographical division of labor to the means of transportation. The small volume of internal commerce in the years before 1800 points to self-sufficiency, at least so far as the products of other states were concerned. This was in general true. The commodities exchanged between the states were relatively few. The northern states needed some of the products of the southern states such as rice and tobacco, and Medford rum found a ready market in the other states, but so far as the actual means of subsistence was concerned the individual states were fairly well able to supply their own needs. They did however need the manufactured goods of England, the products of the West Indies, and the goods of the tropics and the East which the merchant sailing the China seas could procure. The surplus production of the states was chiefly useful in securing these wares.

This situation was not a stable one. We have already noted in another place the stimulus which manufacturing received in the first quarter of the nineteenth century from the wars and from the tariff barriers erected subsequently. A profound change was also taking place in the South. In 1793 the invention of the cotton gin by Eli Whitney had made the growing of cotton commercially profitable, and cotton culture, starting in the states on the eastern seaboard, rapidly spread west and soon became of paramount importance. This provided a new basis of exchange between the North and the South. The New England States bought the southern cotton for use in their own mills, or northern merchants bought it for shipment to Europe while selling manufactured goods of the North or of England to the South.

There was also a marked extension of the cultivation of tobacco in Kentucky and Tennessee. Cotton and tobacco could be rolled from the plantation to a small river and then conveyed by water to one of the important markets such as New Orleans. Both tobacco and cotton could stand the cost of this means of transportation because of the high value in relation to the weight and bulk.

It has frequently been asserted that slave labor, upon which the southern planters relied, was most efficient when applied to one staple crop. But diversification of crops was essential unless an outside supply of food could be assured. The farmers of the Ohio and the Mississippi valleys and of the Northwest provided the foodstuffs at a low cost, and the market thus obtained enabled them to purchase the manufactured goods they required.

We find therefore three main currents of trade. Foodstuffs moved down the Mississippi and were sold at New Orleans or other points, but few goods were shipped up the Mississippi. Cotton and tobacco were shipped north or exported to England. Manufactured goods of British or northern origin came from the North to the southern cities and plantations. On the east and west line goods were shipped in two directions, foodstuffs from the West for export or to supply the growing manufacturing and mercantile population, and manufactured goods sent in exchange by the roads over the mountains or by the Erie Canal.

That this specialization according to the natural resources of the respective regions was mutually advantageous cannot be doubted. The price of flour in Cincinnati rose from three dollars a barrel in 1826 to six dollars a barrel in 1835, and corn rose from twelve cents to thirty-two cents a bushel. Articles that previously had not been marketable could be sent to the eastern markets for sale. The lack of a market for farm products had forced the growth of home manufactures of every sort, but with a rapidly growing market it was cheaper to grow corn or wheat and buy the manufactured goods needed. The cotton culture expanded enormously and, because of the insistent foreign demand, was more profitable than diversified farming could have been. The New England States, New York, and Pennsylvania specialized more and more in manufacturing and by exchanging their manufactured goods obtained their supplies of food more cheaply than by growing them.

Of course it is true that geographical factors called forth the improvements in the transportation facilities which we have noted. The advantages of specialization were so patent that improved means of transportation were inevitable, but once having been established they permitted the process of specialization to continue at a rapid pace.

The beginnings of the railroads, 1830–1850. The canals, important as they were in supplementing the rivers as carriers of the wares of commerce, did not prove a final solution of the transportation problem. The most important rivers ran north and south, and, while the rivers on the Atlantic seaboard ran east and west, they were navigable only to the fall line. In the northern part of the country the canals were frozen during a part of the winter and hence closed to traffic. There was need for some means of transporting people and goods east and west — something whose course would not be determined by the nature of the country and whose utility would not be at the mercy of the climate. This need was met by the steam railroad.

The period from 1830 to 1850 can be described as one of experimentation. The problems of locomotion, of road construction, and of road operation were so novel that it is scarcely a cause for wonder that the first roads do not bear much resemblance to those of today. The pioneer in railroading, the Baltimore & Ohio, first used sails and horses as motive power, but soon after the road was opened in 1830 it tried out and adopted the steam locomotive. The first rails were of wood covered with straps of iron, and these necessarily limited the size and speed of the trains.

There were many other features which seem crude in comparison with present railroads, but perhaps the greatest contrast is to be found in the conception of the possibilities of the railroad. We think of them as vital to the transportation of local and long-distance freight as well as of passengers. In the thirties they were of most importance in carrying passengers; the freight traffic amounted to very little. Most of the roads were local in character, only a few miles in length, and were not connected with each other so as to permit the interchange of rolling stock. The Charleston & Hamburg, one hundred and thirty miles in length, is said to have been the longest railroad in the world in 1833. Even where roads did meet it was usually necessary to unload and reload freight because there was no common width of tracks.

The railroads did not offer the rivers and canals much competition. The rates for the carriage of goods were high, and the service afforded was not superior to that of the waterways. The southern railways were almost the only ones, aside from the coal carriers of Pennsylvania, that definitely anticipated the carrying of freight as an important source of revenue. These were built of uniform gauge to facilitate the transfer of freight cars from one line to another.

It is therefore not surprising to find only about 9,000 miles of track laid in this period. Most of this was in the eastern and southern parts of the country, relatively few miles had been laid in the Middle West, and there were no direct connections between the large seaboard cities of the East and the new cities of the West. A beginning in railroading had been made, but it was not such as to alter the economic life of the people to any marked extent.

The growth of the railways, 1850–1880. The phenomenal development of the American railways came in the second half of the nineteenth century; each decade saw an enormous increase in the number of miles of new track laid. Up to 1848 the yearly average of new track had been between 400 and 500 miles; in 1848 over 1,400 miles of track were laid, and this increased pace was kept up so that by 1860, 30,000 miles of track had been laid. There were several reasons for this. The railroads were emerging from a stage of experimentation, and their possibilities from a commercial standpoint were more clearly understood. More was known about the technique of operation and construction. Iron rails were being substituted for wooden rails, and their use was encouraged by the repeal of the customs duty on foreign rails in 1843 and the beginning of domestic manufacture in 1844. This change made it possible to use heavier cars and engines and promoted economy in operation.

In the fifties connections were made between the eastern and the mid-western cities by the extension of the lines now united into the New York Central Railroad and by the Erie and Pennsylvania Railroads. This bound the East and West together economically, as it provided quick access to the markets of the

East and a ready means of sending manufactured goods to the West, and the merging of economic interest brought about a community of political interest which was destined to bear fruit in the following decade.

Railway building continued in the East and in the South in the fifties, but the most active field of operations was in the states of Ohio, Illinois, and Indiana, where by 1860 nearly 9,000 miles of track had been laid. The eastern states north of the Potomac River had in 1860 about 10,000 miles of railways, and the southern states nearly 9,000 miles. Only a beginning had been made in the states west of the Mississippi, where there was a total of a little over 2,000 miles of track, Missouri leading with 800 miles.

The building of railroads came to a stop with the Civil War. At the end of the war the southern roads were prostrated, and the north and south lines were at a standstill, but road-building in the West went on with undiminished enthusiasm. By 1870 the mileage of the country had almost doubled, and by 1880 it had reached a total of 93,000 miles. Building in fact far outstripped the increase in population; in the decade from 1870 to 1880 the mileage increased two and a half times as fast as the population.

Government investment in railroads. This rapid increase in rail facilities was not entirely dependent on private capital. The various governmental authorities, federal, state, and local, assisted in financing the building of the roads. States made land grants to assist the roads, and the federal government followed the example by direct aid and subsidy. To aid in the building of the Union Pacific for example the federal government made a land grant of ten alternate sections per mile and subscribed $27,000,000 to the purchase of bonds. The land, at the conventional figure of $1.25 an acre, was worth over $14,000,000. In the whole period it is estimated that the federal government gave away 26,000,000 acres or about 40,000 square miles of land, and that the states and the federal government together donated about 242,000 square miles of territory, an area almost as large as the state of Texas and considerably larger than France. It is estimated that the

4) 14 000000
3,500000

investment by all the governmental bodies in stocks and bonds of railroad companies amounted to about $700,000,000.

Private capital did not hold back but was invested freely, for there was unbounded confidence in the future of the railroads, and later on railroad securities offered unlimited possibilities for speculative profits. It is important to note however that the railroads of the United States were not so private an undertaking as is sometimes thought, a fact of some significance in connection with the modern problem of public regulation and modern proposals of government ownership.

Progress in organization, construction, and operation. During all this period there were changes in organization, in construction, and in operation, which were bringing the railroads nearer the present stage of development. The use of steel rails began in the sixties, and, although their use was not widespread until after 1870, their introduction marked a great advance, for they permitted much heavier loads than the iron rails and hence fostered improvements in cars and in locomotives. Independent rail lines were consolidated into long-distance lines. Before 1850 there was no system which had a continuous line of more than 400 or 500 miles. In the period from 1850 to 1870 there were consolidations of local independent roads to form what are now the New York Central, the Pennsylvania, and other systems. This process continued in the following decades until our great railway systems were formed, some of them controlling over 10,000 miles of railway lines.

The formation of through lines brought with it many advantages. For one thing it forced a more uniform gauge, and gradually the present standard gauge of four feet eight and one-half inches was adopted. But, far more important, it encouraged the hauling of long-distance traffic and brought about economies which short independent lines, with frequent loading and unloading of traffic, could never have secured. Four roads — the Pennsylvania, the Baltimore & Ohio, the New York Central, and the Erie — had through connections with Chicago by the seventies. The export trade in grain from Chicago and the regions east of Chicago was increasing in importance, but there was not enough to utilize the carrying capacity of the four roads. Competition set in to secure this traffic, and a rate war of the most bitter sort ensued. The results were important. In the first place rates were permanently lowered; secondly the lowering of rates, in many cases below the cost of transporting the goods, forced economies in operation; and in the third place the roads became definitely the carriers of an increasing volume of the grain trade.

Competition of rail and water routes. Until these combinations took place the railroads did not seriously interfere with the canals, but as soon as they began to bid for traffic over long hauls the canals were doomed. The Erie Canal was one of the few that could withstand the competition of the railroads, and even its traffic declined in relative importance.

In 1865 the Erie Canal carried more ton mileage than the Erie and New York Central railroads combined, but in the next fifteen years the volume of traffic over the canal remained constant while the volume of goods handled by the railroads showed a steady increase. Before 1860 the rates on the railroads rarely fell below two cents a ton mile, but by 1867 they had decreased to 1.9 cents and in 1877 to 1.28 cents.

This process was not confined to the roads in competition with the Erie Canal, but was general throughout the country. Even the Mississippi River began to show the effects of rail competition. After 1880 more and more traffic was diverted to the railroads, although the river always has remained an important artery of commerce.

Railway abuses. The seventies were also notorious for the frauds practiced in the issue and the manipulation of railroad securities, as well as for the favoritism toward individuals and localities and the utter scorn for the public interest or public opinion shown by the railway managers. The abuses which arose in connection with railroads led to a public demand for their regulation either by individual states or by the federal government, and, as we shall see in a later chapter, both the states and the federal government in the next decade enacted more or less stringent regulatory measures.

The growth of the railways, 1880–1920. The railroad mileage continued to grow with extraordinary rapidity. From 1880 to 1890, 70,000 additional miles were built, chiefly in the Far West and in the Southwest. The first transcontinental railroad had been completed in 1869 by the joining of the Union Pacific and Central Pacific railroads at Ogden, Utah, and in 1881 the second of the transcontinental railways was finished with the union of the Southern Pacific and the Atchison, Topeka & Santa Fé railroads at Deming and El Paso. The Northern Pacific was opened in 1883–84, and the following year the Great Northern and the Canadian Pacific lines reached the Pacific Coast.

In the next twenty years (1890–1910) nearly 90,000 miles of new track were built, but after 1910 the pace slackened somewhat, and the World War practically put an end to further building. In 1920 the total mileage in operation in the United States was 259,941 miles; if we include the double tracks and the yard and switching facilities, it amounted to 406,579 miles. The slowing down of the pace of railroad building seems to indicate that the country has been supplied with fairly adequate rail connections and that future development will lie in the direction of building feeders for the main systems and double tracking the lines already in existence.

Economic services. Railroading is one of the most important of all our national industries, both from the nature of the services rendered and from the standpoint of size. The capitalization of the railroads taken as a whole is over $18,000,000,000, and about 2,000,000 men and women are employed on them. In 1920 over a billion and a quarter passengers were carried an average distance of thirty-eight miles, and almost two billion and a quarter tons of freight were hauled an average distance of 181 miles.

Certainly from the standpoint of the railway manager and probably from that of general interest, the freight traffic of the railways is more important than the passenger traffic. For many years the freight receipts have amounted to about seventy per cent of the total receipts. As freight carriers the railroads provide a swift means of abridging distance and are far superior in this respect to the waterways. Goods can be sent when and where they are wanted. This permits the geographical specialization which we have frequently mentioned and furthermore encourages complex division of labor and large scale production. Regions rich in natural agricultural resources can provide the means of subsistence for cities located thousands of miles away. Goods which would perish during the slow journey up a river can be conveyed rapidly and cheaply from the producing districts to great cities. Manufacturing establishments can be concentrated in a given locality and can reap the natural advantages which the region has to offer as well as those which are the result of the propinquity or development of subsidiary industries.

The passenger traffic is nevertheless of great importance both economically and politically. The New England roads for example rely to a much larger extent than the other roads on the income from the normal passenger traffic and the extra volume of traffic during the tourist season. But from a broader point of view the passenger traffic is significant. Anyone who witnesses the daily pouring of hordes of workers into any large city can realize the extent to which the passenger service has made it possible to avoid even greater congestion in the cities than we find. Passenger service has increased the mobility of labor, has made it easier for a laborer to go from one city to another or from one state to another in search of employment, and hence has tended to prevent at the same time a dearth of labor in one section and an oversupply of labor in another. The facility with which communication takes place between the various regions of the country promotes mobility of capital as well as of labor. In the newer parts of a country there is frequently need of a large volume of capital for the development of the resources of the region, and reliance has to be placed on the savings of the older portions of the country. Capitalists are more ready to invest their savings in the industries of a distant region when it is possible by a personal visit to acquire a first-hand knowledge of the nature of the investment and when communication is both swift and sure.

The railroads also render important services in transporting the mails and express. The postal business is entirely a function of the federal government, but the actual transportation is in the hands of the railway companies. Some mail is delivered to them in sacks addressed to a particular place, and mail cars are used in which other mail may be sorted in transit. The express business is likewise in the hands of independent companies, but the railroads operate the express trains and assist in the rapid movement of small packages by attaching the express cars to passenger trains.

This is only the briefest sketch of the organization and operation of our railroad system, but it should suffice to give the reader some general notion of the great services rendered by the railroads and of their strategic position in the modern economic organization.

Traffic on rivers and lakes. At the present time the railroads are the most important means of transporting goods and people within the country. The decline in the importance of the rivers and canals has already been noted. The Erie Canal reached its point of densest traffic in the years from 1886 to 1890 and since then has been declining absolutely and relatively in tonnage carried. The traffic on the Mississippi and Ohio rivers has likewise shown a falling off in the past twenty years.

On the other hand the traffic on the Great Lakes has increased enormously within the last fifty years. While the tonnage of vessels employed on the western rivers was almost cut in half in the period from 1903 to 1920, there was during the same period an increase of over a million gross tons of shipping on the Great Lakes. Immense quantities of ore, coal, grain, lumber, and other commodities are transported annually from the rich regions bordering on the Great Lakes to the eastern manufacturing and consuming centres. The Sault Ste. Marie Canals, which connect Lake Superior with the other lakes, are among the most important traffic channels in the world, and the growth of their traffic may serve as a sort of index to the growth of the traffic of the lakes as a whole. In 1900, 25,600,000 tons of freight were carried through the canals, and by 1920 this amount had more than tripled.

Highways. The highways have been gaining in commercial importance in the last few decades. Always essential in uniting the producing and consuming centres with the shipping centres, the advent of the automobile and particularly of the motor truck has renewed interest in their development. For short hauls, the truck offers very effective competition to the railway, while the automobile has cut into the volume of short distance passenger traffic over both the electric and the steam railways.

Each year large sums of money are appropriated by the federal, state, and local governments to be spent in improving the roads, and each year sees an increase in the total road mileage and in the proportion of surfaced roads to dirt roads. In 1919 there were 2,478,000 miles of public roads in the United States, of which twelve per cent were surfaced. Massachusetts had the largest proportion of roads which had been surfaced (forty-eight per cent), and Ohio the largest number of miles of surfaced roads. During the same year over $200,000,000 were spent on building or improving roads; in Pennsylvania alone over $20,000,000 were spent. An expansion of the work of road-building with particular attention to the type of road which can survive under heavy trucking can certainly be expected in the future, and probably with the increase in surfaced roads will come an enlargement of the area which can be served by trucks.

The merchant marine. We have already noted the place which the merchant marine occupied in the first half of the nineteenth century. In 1800 the tonnage employed in the foreign trade and the whale fisheries amounted to 670,000 gross tons, and in that year American ships carried eighty-nine per cent of the imports and exports of the United States. The European wars gave the American carrying trade a tremendous impetus; during the next ten years over 300,000 gross tons were added to our merchant fleet, and in 1810 our ships carried ninety-two per cent of our exports and imports, the highest percentage up to or since that year.

In the years following there was a decrease in the size of the merchant marine, due in the main to the reëstablishment of peace abroad and the resumption of the carrying trade by the warring nations. Yet our shipping flourished and for the half century as a whole showed a considerable increase. By 1860 the tonnage of vessels engaged in foreign trade amounted to about two and one half million tons, and during that year sixty-six per cent of our foreign trade was carried in vessels of American registry.

The period from the opening of the Civil War to the opening of the World War was one of steady decline in the importance of the merchant marine. In 1910 we had only about 100,000 more tons engaged in foreign trade than in 1800, and American shipping carried only nine per cent of our foreign trade.

Our predominance in this branch of transportation in the first half of the century was due to two things, first the presence of a plentiful supply of materials for building ships together with shipbuilders who could more than compete with European shipbuilders and secondly a body of sailors which was unsurpassed in capacity and daring. The lack of opportunity on land attracted large numbers of the youth of New England to the water. Many of them had an interest in the ship's cargo and therefore felt as much of a sense of proprietorship in the venture and interest in the ship's welfare as the captain himself, who was frequently the owner of the ship and of a large part of the cargo.

The day of the wooden sailing vessel was over when once the practicability of the iron and then of the steel steamer was demonstrated. The deposits of coal and iron and the earlier development of her metallurgical industries gave England an advantage in the making of iron and steel vessels, and the shipbuilders of the United States were slow in following her lead, as can be seen by the fact that the tonnage of sailing vessels of American registry engaged in foreign trade exceeded that of steamers until the year 1902. A part of the merchant marine was sunk or sold into foreign registry during the Civil War, and the high tariffs on shipbuilding materials after the war discouraged the replacement of the vessels lost. But perhaps most important of all, the opportunities of the West presented a stronger appeal than the opportunities of the seas. The United States seemed unable to build and operate ships cheaply enough to compete with foreign nations, particularly with England, but had she been able to compete on even terms it is doubtful if capital and labor could have found there such profitable employment as it found in opening the vast resources of the new states.

The World War brought about a change in the American merchant marine which may be destined to have lasting effects, although it is too early yet to make any predictions with safety. The imperative need of our allies for foodstuffs and munitions stimulated the building of ships in America even before the United States entered the war, and when we ourselves became involved in the struggle one of the first steps taken was to stimulate shipbuilding by every possible means.

For this purpose the United States Shipping Board was established, and a corporation, the Shipping Board Emergency Fleet Corporation, was created, whose duty it was to provide the government with an adequate fleet. Contracts were made with private shipbuilders, and the corporation entered upon an extensive building program of its own with such success that when building operations were completed shortly after the close of the war the tonnage of the American merchant marine engaged in foreign trade had increased to over 11,000,000 gross tons and the American merchant marine occupied second place among the merchant fleets of the world.

The history of the merchant marine since the war has given evidence of the inherent weakness of the American merchant marine. The government's shipping operations have been conducted at heavy loss to the taxpayers, and great difficulty has been encountered in the attempt to dispose of the government ships to private owners. Higher building and operating costs seem still to hamper our competitive power. The subsidizing of the merchant marine has been proposed on the grounds either that it would enable it to get on its feet and become self-supporting or that we need a merchant marine in times of national emergency and consequently it matters little whether it be self-supporting or not. The economic principles involved in this problem will have our attention in a later chapter.

Coastwise shipping. The coastwise trade has always been in American hands, as our laws have forbidden the trade to vessels of foreign registry and further have prevented the admittance of foreign-built vessels to American registry.[1] In 1800 the tonnage of the vessels in the coasting trade amounted to about a third of that engaged in foreign trade, but, as domestic commerce increased, coastwise trading assumed a new importance, and the merchant marine grew in response to the demands placed upon it. In 1850 there were 1,400,000 tons of shipping engaged in the Atlantic and Gulf coast trade, and in 1852 it is estimated that the value of the cargo handled by vessels along the seacoast amounted to about $2,600,000,000. The coasting trade was the most important part of internal commerce and was estimated to have amounted to six times the value of the foreign trade.

Under the influence of railway competition the coasting vessels have lost relatively, but they are still important in shipping bulky goods which cannot stand a heavy transportation charge and which need not be shipped by a more rapid method. In 1920 the total tonnage of all vessels in the coasting trade amounted to 6,357,000 gross tons, including in this 120,230 tons employed on the western rivers and 3,139,000 tons on the Great Lakes. There seems no reason to anticipate any great change in the relative importance of the coastwise trade, unless there are significant changes in railway transportation charges which might divert traffic either to or from the coasting vessel.

---

[1] During the World War these restrictions were relaxed, but they are again operative.

CHAPTER X

THE ECONOMIC FUNCTIONS OF GOVERNMENT

Government and the economic organization. No account of the division of labor and no picture of the economic organization of modern times, or of past times for that matter, would be complete without attention to the part played by the government[1] in the production, distribution, and consumption of wealth. The existence of government is familiar to everyone, and virtually everyone recognizes the necessity of government, the real anarchists being numerically insignificant. But it is to be doubted if many persons have ever stopped to form a mental picture of the economic functions of government or to appraise the important place of the government in the modern economic organization.

Consumption of wealth. Economic activity is devoted to the production of goods and services which are consumed in the satisfaction of human wants. In modern times the greatest consumer of goods and services is the government. In 1923 the economic activity of the people of the United States brought forth a total income which is estimated as worth sixty-seven billion dollars. Of this total national income the government (including all grades, federal, state, and local) obtained for its own uses ten billions. It is true that a good deal of this was passed on by government to various citizens in the form of interest payments, amortization of debt, pensions, etc. Yet the fact remains that more than one seventh of the total national income was, in the first instance at least, claimed by the government.

---

[1] The term "government" is used in this book to include all grades of government of any particular community, for example in the United States of America the federal government, the state governments, and the governments of counties, towns or townships, cities, school districts, and all other political jurisdictions. It is rather common in popular speaking and writing to refer to the United States federal government as "the government" in distinction from the states and the local jurisdictions. The term is not so used in this book.

Production of wealth. On the side of production, the government is ever present. For example, the United States Post Office furnished services in the fiscal year ending June 30, 1925, which cost about six hundred and forty million dollars. The United States Government Printing Office is the largest printing establishment in the world. Our state governments are heavily engaged in the building of highways, cities furnish their citizens with water, street lights, street railway service, etc., and there are numerous other examples, which will occur readily enough to the reader, of commodities and services produced by national, state, and local governments. But after all direct production of this sort represents only the minor part of the activities of government. The chief contribution of government to production is indirect, consisting in such services as protection from outside foes, enforcement of law and order among the people, promotion of the public health, public education, etc. These services, though in a sense indirect, are absolutely essential to production, and they impinge upon every department of national production. Being essential services, which must be performed by some agency if production is to flourish, their performance by government represents a true and important case of the division of labor.

Distribution of wealth. In the field of distribution, the government is no less omnipresent and important. We need only to recall that the various governments in the United States in 1923 took from the people in taxes seven and three quarters billions of dollars out of the total national income of sixty-seven billions. Two and a quarter billions more were taken from the people by other means than taxation, in order to provide the ten billions of government expenditures in that year. The federal government alone paid out about a billion and a half to holders of government bonds in interest and repayment of principal and distributed nearly three quarters of a billion more in pensions. Billions were paid in salaries and wages to nearly three million government employees, comprising more than one ninth of all the persons in the country employed for wages or salaries and over six per cent of all persons in gainful occupations (*i.e.*, including those working for themselves).[1] The government is therefore the greatest payer of interest, the chief disburser of pensions, and the leading wage payer of the United States. Taking and redistributing to the people more than one seventh of their total annual income, as the government does, cannot be accomplished without the most profound effects upon the distribution of wealth and income.

Varieties of government. Far back in the obscure beginnings of human existence, whatever and wherever they may have been, there was the beginning of government. The human race has always been socially inclined, living its life not as isolated individuals, but in organized groups — families, tribes, clans, cities, and nations. For such organized existence there has had to be some authority to control the activities of the individual members for the common good. At different epochs in the history of mankind various kinds of government have appeared, passed across the stage, and made their exit. At the present time many forms of government are to be found in different parts of the world, from the rule of the tribal chief or " medicine man " to the parliamentary government of Great Britain or the Soviet rule of Russia. To trace the historical development of government or even to examine the various types of government existent today would take us too far afield. Our purpose will be sufficiently served by confining our attention to the modern democratic form of government as represented, with innumerable variations of course, in most of the leading nations of the present day. For obvious reasons, we shall find it desirable and convenient to draw most of our illustrative materials from the government of the United States.

Political organization in the United States. In so doing we must not forget that the American government is in certain ways peculiar. Whereas in most countries the national government is the sole fountain head of sovereign power, every American schoolboy knows that the United States is a federation, in which sovereignty is divided between the national government and the

---

[1] The figures for 1920, the latest census year, are:

Total persons ten years of age and over in gainful occupations .	41,614,248
Total employees	23,000,000
Total government employees	2,700,000

several states. We are accustomed to speak of three grades of government, (1) national or federal, (2) state, and (3) local. The federal government was formed in 1789 after the people had adopted the Constitution of the United States. Before that there had been thirteen states, which had just won their independence from Great Britain and which were therefore separate sovereign powers. These independent states voluntarily surrendered some of their sovereignty in order to form the national government, which government has only such powers as are to be found either granted to it in the Constitution of the United States or implied as corollaries of the powers so granted. The states, increased now to forty-eight in number, are the original and residual possessors of sovereignty, having all sovereign powers except such as have been granted to the national government. The third grade of government has no sovereignty, but consists of administrative units, counties, towns or townships, cities, boroughs, school districts, road districts, fire districts, irrigation districts, etc., created by the states and with only such powers as have been granted them by their respective states and which may be taken away at the will of the states.

The foregoing summary of the political organization of the United States very probably contains nothing that is new to the reader. It is merely a reminder of facts which are vital to the subject we are about to study. We are here on the threshold of a department of economics which approaches very closely the domains of political science and law, and a clear understanding of the relations between the different grades of government in the United States will be found indispensable.

The economics of government. We may now proceed to analyze the economic activities of government. So intimately is the government related to every branch of the economic life of the people, that before our study is concluded it will appear that the economic activities of government include all governmental activities. Nothing that the government does is without influence upon the production, distribution, and consumption of wealth. In this sense there is no possible separation between political science and economics. The distinction between the two fields lies rather in the point of view and purpose of study. What here concerns the student of economics is the economic bearing of governmental activities; *i.e.*, the relation of government to the production, distribution, and consumption of wealth. It is from this viewpoint that the reader is invited to pass in review the principal functions of the modern democratic government, with special reference to the situation in the United States of America.

Defense. From time immemorial, the human race has found it impossible to get on without conflict among its members, either individually or by groups. Some students of psychology believe that there is inherent in human nature an instinct of "pugnacity" and an instinct of "self-assertiveness" which suffice to explain the constant presence of conflict. The economist need not concern himself with the dispute over the subject of human instincts. It is sufficient for him to recognize the fact that a certain amount of fighting appears to be the normal thing among human beings. Men fight in defense against interference with their activities or wishes, as a protest against insult, or sometimes even for the sheer joy of combat. Groups of men — families, tribes, or nations — engage in conflict with other groups, either aggressively, to seize the possessions of others, or defensively, to protect their own possessions and interests. Fundamental to such group conflicts is the perpetual struggle for existence, through which each group is impelled by the motives of self-preservation and self-perpetuation to seek for itself the greatest possible share of the limited products which nature offers to mankind.

Now this tendency to conflict necessarily leads to organization and authority, and from two different directions. Obviously the group cannot wage successful warfare against other groups without organization and some governing authority with power to marshal and direct the group for the purpose of attack or defense. Thus the first and most important of the functions of government is defense, by which we mean defense of the group against attack by outside foes, protection of the rights of the citizens against foreign nations and their citizens, and also (though possibly here we stretch the term "defense") aggressive attack against other nations. In the modern nation this function demands much of the activity of the general government and specifically calls for the services of army, navy, air defense, consular and diplomatic services, etc.

The economic significance of this function of government will be obvious as soon as it is regarded as an exemplification of the division of labor. Production cannot proceed successfully if the nation's territory is frequently overrun by foreign foes, or even if the danger of such invasion is ever imminent. One of the fundamental causes that enabled England to forge so far ahead of the Continental peoples in economic development was that her insular position protected her from attack and made it possible to develop her sheep and cattle, her agriculture, and later her trade and industry, at a time when the lack of security was a fatal obstacle on the Continent. Defense against outside foes must be organized, and to be efficient it must be specialized. The latter requirement was one of the principal causes of the breakdown of the feudal system. The king found that his armies, hastily assembled from the agricultural population on each occasion of conflict, made an inefficient fighting machine, while at the same time the resulting disorganization of the nation's productive forces undermined his military strength through failure of supplies at the very time when they were most urgently needed. Hence the commutation of the feudal obligation of military service into regular money payments, thus providing the funds for organization and maintenance of a standing army. Production was promoted by the presence of a superior defensive force and by relieving (to a great degree at least) the agricultural and industrial population from interference. In like manner the early American settlers found it necessary to post armed sentries about their fields to guard against attack by the Indians and to enable the agricultural workers to pursue their labors undisturbed. These are examples of division of labor in production, and the student will not fail to recognize that, wisely employed, the army, the navy, the diplomatic service, and the other governmental agencies of defense are just as properly included among the nation's productive forces as is any other group of the people.

Justice and security. It was stated above that the pugnacious tendency of mankind leads to organization and authority from two different directions. From the first direction we get the defense function of government. From the other direction comes the function of justice and security. Conflicts arise, not only between different groups or nations, but among the individual members of a single nation. In the earliest days of any community the people settled their private disputes among themselves. The comparatively modern history of the frontier in America illustrates this phase. The colonial settlements upon the Atlantic seaboard either arrived with some form of government already organized or established one immediately. But as the more adventurous ones pushed on to the West and established successive waves of frontier communities from the Alleghenies to the Pacific coast there came in each community a period in which there was little in the way of effective governmental authority. Each ranchman or miner had to go armed and be prepared personally to defend his property and person. Disputes, whether over mining claims or cards or women, were settled by private fights or attack from ambush. A state of lawlessness and of nervous fear and tension was normal.

Now it is obvious enough that production, which requires for its successful promotion security of life and property and a state of order and tranquillity, was hindered thereby. There arose therefore first the vigilance committee, by which the people sought to act as a group and more or less authoritatively against the horse thief, the claim-jumper, and other common foes. The vigilance committee was a makeshift government pressed into service to maintain a certain semblance of law and order until such time as the duly constituted government was prepared to take over the service. In modern advanced communities it is fully recognized (1) that private rights must be protected by government in order that the people may be free to devote themselves to their productive occupations undisturbed by the fear of attack or the necessity of being constantly prepared to defend themselves, (2) that government must provide agencies for settling private disputes in order that justice may be done and that the peace and tranquillity of the community be not disturbed by private brawls and violence, and (3) that contracts entered into between the citizens must be enforced by government in order that production and peaceful living be not deprived of the inestimable advantage of contracts freely undertaken and strictly enforced. Thus we have the police, the courts, and various sorts of penal and correctional institutions. This also represents an example of the division of labor, under which the government undertakes a certain function which is thus performed more efficiently and more economically than would be possible otherwise. Production is advanced, and the police, the courts, and the penal institutions, with the persons engaged in their service, are productive agencies as truly as any other occupational class of the community.

The function of defense is naturally entrusted to the national government, in the United States to the federal government. The state and local governments are specifically relieved of obligation for the defense of their citizens against foreign foes, although a certain contribution may be made by the states in aid of national defense in time of war. Justice and security are on the other hand mainly in the province of the subordinate units, the states and local governments in the United States, though the federal government is not entirely devoid of obligation in the service of courts, police, and prisons.

Regulation of industry. In an earlier chapter of this book there was discussion of the mechanism which controls the direction of industry under the prevailing system of division of labor. It was there pointed out how in its first stages industry was held under the control of law and custom, till finally it broke through the shackles that were impeding its progress and gained freedom under the principle of *laissez faire.* Competition, acting through the instrumentality of price, was entrusted with the task of giving wise direction to industry and protecting the interests of both producers and consumers. In a still later stage it turned out that competition was not quite up to the task. Today therefore while in general competition controls industry, there are various points at which return to some degree of control by legal authority has occurred.

This is to be found principally in two kinds of industry: (1) those that tend naturally to become monopolies and (2) those which are so complicated in their organization and operation as to be only imperfectly understood by the public. As examples of the first we note the railroads, the street railways, the concerns which furnish water, gas, electricity, heat and power, etc. The monopolistic character of such enterprises must be accepted and, in place of competition, the government must undertake to regulate the character of the services rendered and the prices charged. The United States federal government, principally through the Interstate Commerce Commission, and the state governments, through their railroad and public utility commissions and similar agencies, exercise this regulation. Examples of the second class of industries requiring legal regulation are banking, insurance, etc. Here the complications of the businesses are such that the public is not able to protect itself relying only upon competition. So most of our states have banking commissions and insurance commissions to regulate these respective enterprises. The federal government exerts a certain control over many kinds of business through the Federal Trade Commission. Cities frequently exercise similar legal control over taxicab companies, peddlers, etc.

The growth of monopoly and the technical complications of modern industry put the consumers at a disadvantage as against the producers and even give certain producers advantages over others. Government then steps in and lays down rules for the equitable conduct of business and acts as umpire for the enforcement of the rules. This is an economic service of the utmost importance. It facilitates production by limiting unfair practices among producers and by enforcing standards of product or service and reasonable prices. The effect upon distribution and consumption is obvious. This function of government will be discussed at greater length in a subsequent chapter.

Protection against disease and accident. Ill health, disease, personal injury by accident, and death from disease and accident are dangers which everyone has to face. Most persons exercise some degree of care to protect themselves, and until modern times it was not supposed to be the function of government to protect the people against these dangers. The true nature of disease and the means by which it is carried from one person to another and from one region to another were imperfectly or not at all understood. Disease, accident, and death were regarded as " acts of God " to be accepted as they came with fortitude and a submissive spirit; that organized effort could do anything to prevent or mitigate them was not realized. The marvellous discoveries of modern science and the equally marvellous accomplishments of medical practice have opened up a vast field for profitable government action. The causes of many diseases may be eliminated, and the spread of disease may be checked. Certain of the most deadly of diseases, such as yellow fever, smallpox, typhoid, etc., have been virtually abolished in certain parts of the world, and their ultimate complete disappearance is within the limits of possibility. Modern treatment of disease is able to limit its power and thus reduce the mortality and the permanent after-effects of disease. These are results however which cannot be accomplished by individual action, no matter how intelligent and courageous. The enlightened action of the few will be neutralized by the ignorant and selfish action of the many. Results may be accomplished only when all are brought under the compulsion of government authority.

To enumerate and describe all that modern government does in this field would be too long a story for this book and take us too far afield from the object of our study. Let it suffice merely to call to mind some of the things that are being done. Prevention of disease is the aim of modern sanitary regulations, which seek to ensure pure water supply, to prevent the sale of unwholesome milk, meat, fruit, and vegetables, to enforce proper sewage disposal, etc. Living conditions in cities are regulated by tenement house laws setting certain standards of light, air, and space.

The spread of contagious diseases is checked by compulsory reporting of cases and by isolation and quarantine imposed by the local public health authorities. The federal government through its quarantine and immigration services exercises constant care to prevent the entry into the country of diseased persons. Certain especially dangerous trades, such as the manufacture of matches, are regulated or have been entirely reorganized under rules imposed by governmental authority. The treatment of disease is no longer left entirely to the resources of the sick themselves or their families and friends. Public hospitals, clinics, and dispensaries come to the aid of the people, ~~especially~~ those too poor to provide ~~proper treatment at their own cost~~.

On the side of accident prevention the government is no less active. The law enforces certain standards in the construction of factories, stores, school buildings, theatres, and in general all buildings where large companies of people assemble. Fire exits and fire escapes must be provided; stairs must be lighted and kept free from obstruction; stairways, elevator shafts, etc., must be properly guarded; dangerous machinery must be screened; these are only a few everyday examples of the activity of government in the effort to prevent accident.

The practice of medicine, dentistry, and even hair-cutting is generally restricted by the imposition of certain standards of education and experience and by sanitary rules and regulations.

In all this expansion of government activity the necessity of individual care of health and life is of course not abrogated. Indeed much of the organized work in the interest of public health is carried on by private wealth and initiative. Magnificent hospitals and dispensaries are frequently privately endowed, and private charity does much for the indigent sick. The scientific study of disease and the practical efforts to stamp out the great epidemic scourges have received much of their impetus and support from private generosity. Witness the magnificent world campaigns so successfully waged by ~~the Rockefeller~~ Foundation against yellow fever, malaria, hookworm, etc. Government does not at all supersede private effort. But it adds resources of great importance, and when compulsion is necessary the government brings to bear the indispensable force of authority which it alone possesses.

The economic bearing of the sanitary function of government is evident. Of the three factors of production ordinarily mentioned (land, labor, and capital) man himself is one. All production depends upon him, and nothing is more essential to his success in production than physical health and vigor. As an individual, the sick man is inefficient or useless or worse. On the other hand whole communities, states or nations, have had their productive powers impaired by such scourges as the "black death" of medieval England, the yellow fever of Panama, the malaria and hookworm of some of our southern states. Government, in what it does to prevent and cure disease and to remove the causes of accident, is contributing powerfully to the production of wealth, promoting its equitable distribution, and adding to the enjoyment of its consumption.

Care of dependents and defectives. In every group of people there will be found certain individuals who for one reason or another are unable to provide for their own needs or to contribute their share toward the common provision. The cause may be some congenital physical or mental defect, or the result of sickness or accident, or even some abnormal twist of character which incapacitates the individual as a worker. Among primitive and savage peoples, such individuals were usually recognized as a burden and handicap to the tribe; they were consequently left to perish of want or were even executed on the authority of the chief. Modern sensibilities cannot tolerate such a summary, though logical, solution of the problem. We feel that, for our own peace of mind if for no other reason, the poor, the sick, the crippled, the insane, and the other unfortunates must be cared for and made at least tolerably comfortable. Of course many such are cared for by their families and friends and so do not present a public problem. Others are supported by private charity. In medieval times the Church was the principal agency for the care of the large classes of dependents and defectives which were generally to be found in every community. As the Church lost its authority and its wealth, it gradually ceased to perform this function. Private charity was more insistently called upon, and finally the government was given the task of providing for all those who were not otherwise cared for. In general the government in this undertaking has proved more efficient and less demoralizing to the dependents than either the Church or private charity. The government alone is able to exert the comprehensive authority which this function demands for its effective accomplishment, and only the government is able to distribute the burden of cost equitably among the self-reliant parts of the community. This is not at all to belittle the accomplishments of private charity when properly organized, as it is in many of the larger cities, or to underestimate the large part of the burden of which the government is thereby relieved.

This function is in the United States performed principally by the state and local governments. The states generally provide the insane asylums and certain hospitals; the poor-house is usually a county or town institution; the large cities are active in poor relief and in provision of hospitals and other agencies for the care of the sick and crippled.

The economic aspect of this function presents peculiarities and some interesting problems. So far as the mere production of wealth is concerned, this function is generally a handicap rather than a help. The dependent and defective classes are in the nature of the case not productive, and their support is a burden upon the productive members of the community. From this point of view the summary method of the savage tribe is the more logical one. However modern civilized man insists that the dependent classes be cared for, and he is willing to pay the cost. The economic result is that a certain part of the income of the community is diverted from those to whom it would otherwise go in order to care for those who are unable to provide for themselves. Production is not enhanced, may indeed be hindered, while a material change is effected in the distribution and consumption of the product. It is an interesting question how far a nation may go in this direction without threatening its future in the struggle for existence. If the dependent and defective classes are supported in ease and comfort, if they are permitted to perpetuate themselves — and many of the traits which characterize them appear to be hereditary — the time may come when their numbers have become so great and the burden of their support so heavy that the capable members of the nation will falter under the load and succumb in the struggle for existence with other more ruthless nations. Or the diffusion of the inferior traits of the dependent class among the whole people may so weaken the national stock as again to handicap it in the struggle for existence. It must be admitted that our modern attitude is directly contrary to the evolutionary principle of survival of the fittest. We permit survival of the unfittest, particularly if the dependent classes are left free to marry and have children.

Of course it is not intended to suggest a reversion to the ruthless methods of the savage tribe. Nor is this the place for an exhaustive study of this subject or the formulation of rules whereby the modern humanitarian sensibilities may be respected without danger to the race. Our purpose is served when it is clearly seen that here is an important function of modern government, with immediate effects upon the production, distribution, and consumption of wealth, and whose remote effects deserve the most serious consideration.

Maintenance of moral standards. Every group of people has certain approved modes of conduct, the proper ways of doing things, the *mores*, which exercise a powerful control over every department of human conduct. The particular acts which are approved or disapproved differ as widely as the poles among different peoples. With certain races monogamy is the rule; among others polygamy is the correct thing. The taking of human life is against public morality among most civilized peoples; but with certain less advanced tribes killing, provided of course proper discretion is shown in the choice of the victim, is highly meritorious or even a sacred duty. But, however much moral codes may differ, however difficult it may appear to find rational grounds for the mores, each community takes its own approved customs very seriously. Those who act differently are at best bizarre and "queer," at worst they are "immoral" or "wicked."

Sooner or later the more important of the mores tend to become standardized and formulated in laws and enforced by the government. Thus modern governments regulate the traffic in intoxicating liquors and narcotic drugs. The laws of nearly every nation take cognizance of prostitution and seek to regulate if not to suppress it. The conduct of people in dance halls, theatres, and other amusement places is subject to legal regulation. Censorship of books, plays, moving pictures, and works of art seeks to prevent the obscene and indecent, meaning that which is flagrantly at variance with the accepted moral standard of the particular group. Dress has always been a subject for the lawmaker and the ruler. The Immigration Service of the United States restricts the entry of foreigners, both in order to keep out those whose conduct is likely to offend our moral standards and also in order to restrict the numbers of the population so as to maintain our economic and social standards against the inrush of alien customs. The economic significance of the mores can scarcely be exaggerated, being as they are an essential part of the environment in which man lives and which limits his economic activities. They affect profoundly the production, distribution, and consumption of wealth. The government's action in formulating and enforcing social standards is correspondingly important from the viewpoint of economics.

Protection against certain forces of nature. Nowhere is the effectiveness of group action under governmental authority more clearly shown, though in a relatively minor sphere, than in the measures taken to give protection against certain destructive forces of nature, such as fire, flood, and storms at sea. The city fire department, the United States Life Saving Service, the state forest fire patrol, the government levees along the river banks are well-known examples of a function of great economic importance which could be performed only imperfectly or which in many instances would not be performed at all if left to private initiative.

Protective and developmental functions. The functions of government to which attention has thus far been invited are alike in that their purpose is protective; that is, to prevent or overcome that which is regarded as evil. They are sometimes called the "protective functions" of government. We must notice now certain governmental functions of a different sort, which aim, not primarily to defend from evil, but rather to promote that which is held to be good. They have been called the "developmental functions" of government. Their nature will appear more clearly as we proceed to the study of the particular governmental activities which are so classified.

Education. Foremost among them comes education. It was not so long ago that there was virtually no public education. Such education as existed was mostly in the hands of the Church. Only the favored few, kings, nobles, and priests, received any education; the masses of the people went on from generation to generation in ignorance. With the decline of the Church's influence, education was gradually taken over both by private schools and by the government. Today universal education, through the elementary grades at least, is the ideal and to a considerable extent the accomplished achievement of enlightened democratic nations. The responsibility for this accomplishment is divided between governmental and private educational institutions. In the United States, as everyone knows, primary education is furnished free, or practically so, by the schools of the towns or counties. High schools are provided by the counties and the more populous towns and cities. State laws generally make school attendance compulsory up to a certain age, and the government is therefore under the obligation of furnishing at least the elementary schools, accessible to all the people. High school attendance is not generally compulsory, yet the facilities of the public high schools are available to the people in almost every part of the land. Even the higher education of the college and the specialized advanced and technical training of the university are provided by state universities in the majority of the states.

Private schools in America coöperate with the government, sparingly in the elementary grades, more generously in the "preparatory" schools of high school grade, and with extraordinary liberality and efficiency in the furnishing of facilities for college and university education, so much so that many of the states, especially in the northeastern part of the country, are able to rely upon privately endowed colleges and universities and are freed from the necessity of providing state universities.

The fundamental motive of universal public education is probably political rather than economic; we believe that democratic government can flourish and survive only when there is a fair degree of intelligence and knowledge among the people. Nevertheless the economic effects of public education are scarcely capable of exaggeration. As producers, as claimants of their respective shares of the product, and as consumers of wealth, the whole character of a people is governed by the degree and the nature of the education which it enjoys. The economic superiority of the United States and the other leading nations of the world, as equally the economic inferiority of such nations as China and India, is without doubt largely to be ascribed to differences in education. With such understanding as the reader already possesses of the factors of production and the principles which govern the distribution and consumption of wealth, the immense significance of education in the realm of economics will scarcely need further demonstration.

Before leaving the subject of education, mention at least should be made of certain governmental agencies, not schools, but having a real educational tendency, such as public libraries, museums, art galleries, industrial exhibitions, and the like. Here again there is material aid from privately supported institutions.

Somewhat further removed from education strictly defined, yet partaking of the same general nature, we find governmental establishments for public recreation. There are parks, zoölogical and botanical gardens, bathing places on seashore, lake, or river, athletic fields, vacation camps, and playgrounds. Modern governments are giving considerable attention to this function, whose economic aspects will be obvious.

Religion. Many nations recognize and support a state Church. The political and social questions thus raised are of great importance but lie outside the field of our investigation. Economically the support of the established Church may be a considerable burden upon the people, which is further increased when the Church is the owner of large estates which may be either withdrawn from the most productive uses or managed solely in the interest of the Church. To the American student this function of government is of less interest, since our traditional policy of separation between church and state has enabled our government to avoid entirely this function.

Facilities for national industry. There are certain points at which the government, on account of its relationship to all the people and its possession of the power of compulsion, is able to furnish facilities for the aid of industry which private business agencies could provide for themselves only imperfectly or not at all. Highways and bridges, formerly built and operated as private enterprises, are now generally furnished by the government for the free use of all the people. The maintenance of harbors and navigable streams is almost universally a government function, as is the protection of navigation by lighthouses, buoys, etc. The government thus enters very directly and helpfully into the production of wealth and services. Production is further aided and the distribution of incomes materially affected by the system of patents and copyrights enforced by government. To government is entrusted the task of furnishing the people their monetary system, or at least their standard money. Forestry and irrigation projects are commonly undertaken by the government in the interest of more efficient production. Determination of standards of weights and measures and the enforcement of their honest use is an important economic function of government. The consular and diplomatic agents perform valuable services to the country's merchants, bankers, and others engaged in foreign trade. There is often also aid to domestic trade, through standardizing patterns and specifications, dispensing crop and market information, etc. Agriculture is aided by state experiment stations. Various official agencies conduct investigations and furnish useful information to the people, as for example the United States Census. A comprehensive list of such services of government would be a long one; enough examples have been cited to demonstrate the magnitude and variety and importance of such contributions of government to the economic endeavors of the people.

The primary functions of government. The governmental functions thus far studied are in general devoted to rendering services which the people, acting as individuals or through private associations, would find it difficult or impossible to perform so well for themselves. This statement, though it would be disputed by the anarchist and by some extreme individualists, represents fairly the consensus of opinion of the great majority of the people of modern civilized nations. It is recognized that the performance of these services requires the power of compulsion and the all-inclusive jurisdiction which government alone possesses. Very few persons would today seriously suggest that any material part of the functions thus far mentioned in this chapter (with the exception of religion, perhaps) should be taken away from government and relegated to private initiative. It is in this sense that these are peculiarly governmental functions, for which we suggest the term *primary functions* of government.

Government industries. The foregoing are to be contrasted with government activities in the industrial field, in which private business occupies the dominant position — activities which are performed or might conceivably be performed just as well or better by private enterprise. Here there is no such general agreement as to the proper scope of governmental activity, and there is room for continual debate as to how far the government may advantageously go in the industrial field. Practically every nation has given over to its government the business of carrying the mails. Many European governments operate also the railroads, the express business, the telegraph, and the telephone. Local governments, in America and still more frequently in Europe, engage in the business of furnishing their citizens with water and, somewhat less commonly, with gas, electricity, street railways, etc. State and municipal forests are common in Europe. In general the expansion of government into the industrial field has gone much farther in Europe than in America. The industrial activities of government must obviously have most important economic aspects and quite unlike those presented by the primary functions. Indeed this whole subject is so important that its treatment will be reserved for a separate chapter in connection with the study of public finance.

Taking stock of our economic organization. We have now arrived at the conclusion of the first part of our study. At the very outset the reader was reminded that man differs from the lower animals fundamentally in the infinite number and variety of his wants and in the extraordinary range and intensity of his activities in the quest of their satisfaction, and Part I of this book has sought to present a picture of the economic organization which has been built up for this purpose. The central theme, giving coherence and meaning to the infinitely complex picture, is that modern economic life is coöperative, based upon division of labor, individual freedom, private property, and the control of production and distribution by price. Through examination of earlier economic systems of a different sort and by the aid of a brief glance at the evolutionary process by which the present system has developed, it has become evident that the latter is more effective in the production of wealth and services for the satisfaction of human wants than any system which has preceded it in the history of mankind.

It must not be presumed however that the productive system works perfectly. No human institution is perfect. The critical student of current affairs has frequent occasion to observe that the productive machine does not always run with perfect smoothness. He notes an occasional breakdown in the transportation system, disturbances in the mining industry which interrupt the normal flow of coal from the mines to the consumers, strikes and lock-outs in the factories, unwise distribution of land between the various farm crops, and so on. It is evident that to many a person the alleged freedom to choose an occupation, to produce what he wishes, and to have and enjoy the product of his industry seems more imaginary than real. It is certain that not all the men whose judgment guides the industrial machine are the ones best fitted for their posts. Capital does not instantly and without exception distribute itself among the several industries in perfect adjustment to their respective needs. There is obvious maladjustment of the labor force, and no one can claim that every parcel of land is always put to its most productive use. Competition appears not always to apportion rewards according to productive contribution but sometimes to favor the merely shrewd or the actually dishonest. Private property, in company with free competitive industry and reënforced by inheritance, occasions, or at any rate permits, great inequality in the distribution of the products of industry and of the capital instruments from which further products are to flow.

It must also be evident to anyone who looks beneath the surface that, when government authority was displaced by free enterprise under the control of price as the directing force back of production, the control of production was actually given over to the consumers, or more specifically to those who have the wherewithal to pay for the things they want. As has been observed, nothing will be produced, no matter how beneficial it might presumably be to mankind, unless people are able and willing to pay a price equal to its cost of production, and any conceivable thing will be produced, no matter how frivolous or even positively harmful, so long as someone is prepared to pay for its cost. Production is thus directed, not toward the maximum of things most beneficial to mankind on the basis of some ethical or moral standard, but toward the things which are wanted by those who can pay for them. It is this which gives special significance to the inequality in the distribution of wealth which appears to be an inevitable concomitant of the present economic system. Those who have abundance do not merely receive a greater share of the product of industry; they have a disproportionate voice in saying what things shall be produced. The machinery of production is directed to turn out steam yachts and race horses, fine silks and costly wines, even though there be those who are in need of the food and clothing which might otherwise be produced in greater abundance; land is set aside for private estates and hunting preserves which might otherwise be devoted to the production of a more abundant flow of food and other necessities.

Shall the system be scrapped? Now there are those to whom these weaknesses of the present industrial régime loom so large that their patience is exhausted and they cry: away with the whole system! Industry is going to rack and ruin in the hands of irresponsible and incompetent entrepreneurs; let us substitute government ownership. Competition is nothing but legalized selfishness, in which the race is to the unscrupulous and grasping, not to the honest and efficient; let us therefore substitute the real coöperation which socialism offers. Worst of all is the intolerable inequality in the distribution of the good things of life; let us therefore prohibit inheritance, let us have government ownership of the capital instruments of production, or better still let us abolish private property entirely and entrust all wealth, along with the direction of industry, to the socialistic state.

What the present system accomplishes. Now the time to pass judgment upon proposals to alter radically the existing economic system is not at the beginning of our study of economics but at its close, when we shall better understand the details of the present system and the possibilities of proposed substitutes. What our present purpose requires is to attain a clear notion of the general outline of the existing system and its operation. The essential foundation of the philosophy of economic freedom, with its corollaries of private property, individual control of industry, and competition, is that, by leaving the several individuals free to act in their own interests, the maximum total production will be obtained and to that extent at least the interests of the whole community thereby served. As was pointed out in the first chapter, the welfare of mankind depends first upon the total amount of production and second upon its distribution. Any " reform " which aims to improve distribution must see to it that the effectiveness of production is not incidentally impaired, lest the remedy be worse than the disease.

There can be no doubt that the system of economic freedom does offer a powerful motive to production and does spur men on to energetic effort. Equally without doubt is it that economic freedom is an incentive to leadership. Executive ability of a high order is by no means over-abundant, and it is absolutely indispensable that such leadership be found out and placed at the disposal of industry. He is most successful who not only perfects the internal organization of his own business but who makes the most perfect adjustment of his business with other types of business round about him. The present system does on the whole call forth efficient leaders and make them coöperate. Again there can be no doubt that economic freedom offers a powerful stimulus to saving. The motives for saving are varied, but they are particularly strong in a system of private property. There is the one who is timid about the future, who fears lest the almshouse or the charity ward of a hospital be his fate in old age; to avoid this he is willing to deny himself things in the present. Another wants his family to occupy a position of equality with those who at present live without work; he labors long and hard, expecting to find his reward in the attainment of that position. Another, not so concerned with social position, is fearful lest those dependent upon him be in want after his death. Other motives might be mentioned, and we would find that most of them are conditioned on the responsibility of the individual for his own present and future welfare, which is a corollary of freedom. The present system thus promotes the accumulation of capital needed for production.

Again it is certain that the automatic control of industry through competitive prices does on the whole compel producers to bring forth the goods the people want in the respective quantities wanted and to deliver them to the consumers at about what it costs to produce them. This is an accomplishment not to be lightly cast aside, even though the direction of industry may not perfectly conform to some preconceived standard of ethics or morals. And finally the present system of distribution, in spite of all exceptions, does bear a very definite relation to the productive contributions of the several individuals.

The present conclusion. It will be noted that the statements of the foregoing paragraph leave ample room for all the exceptions to which attention has previously been called and for the perfectly evident conclusion that the present system falls short of perfection. It is therefore to be presumed that the present system is capable of improvement. In fact the system is already the result of first a reaction against government control and later a material departure from the extreme of *laissez faire.* The Industrial Revolution brought conditions which compelled limitation of the field of free competition in the interests of the laboring class by two types of regulation. In the first place the workers found that they were competing among themselves and that consequently the individual worker was less efficient in bargaining with his employer. So, not without a struggle, the laborers gained the right to combine and to act through their organizations. Voluntary restriction on the individual was agreed upon, and bargaining with the employer by the working group as a whole was substituted. In the second place it was apparent that public regulation was necessary to protect the worker unable or unwilling to protect himself, and the state has passed legislation regarding the hours of work for women and children, the fencing of dangerous machinery, and the conditions of work in mines or in other unhealthy or dangerous occupations. On the other hand it has been found that the entrepreneurs themselves need protection against " unfair " practices developed in their competition with each other. And finally the consumers are today safeguarded by pure food and drug laws, by restrictions upon combinations of producers, and even by legal regulation of prices in certain cases, as for water, gas, electricity, railway transportation, etc.

In such ways has the extreme philosophy of freedom been modified. That further modifications of the present economic system are to come is a reasonable, indeed an inevitable, inference. That the whole system will be scrapped and something quite different substituted is most unlikely, just as unlikely as that it will be crystallized and perpetuated in exactly its present form. We shall give thought to these questions in the concluding chapters of this book. We have now reached the point in our study at which we can at least appreciate the well-nigh infinite complexity of our economic organization and can recognize that he who would attain to such understanding of its nature as to be able to pass wise judgment upon it has set himself no mean task. At the present stage of our study it would therefore be most unscientific either to jump to the conclusion that the imperfections of the present system warrant its discard in favor of something radically different or to conclude from the undoubted services to mankind of the present system that it is perfect and should be defended against any change. In the chapters now before us we shall seek a more complete and intimate acquaintance with the economic system and shall thereby at least approach nearer to that knowledge and understanding which alone can qualify one to sit in judgment upon it.

Suggestions for Further Reading on Part I, the Economic Organization

I. General Economics Texts

Taussig, F. W. *Principles of Economics*, 2 volumes. Third edition. New York, 1921

Ely, R. T. *Outlines of Economics.* Fourth revised edition. New York, 1923

Marshall, A. *Principles of Economics.* Seventh edition. London, 1916

Fisher, I. *Elementary Principles of Economics.* New York, 1912

Shearman, H. P. *Practical Economics.* New York, 1922

Franklin, F. *Plain Talks on Economics.* New York, 1924

Hadley, A. T. *Economics.* New York, 1901

Taylor, F. B. *Principles of Economics.* New York, 1921

Clay, H. *Economics: an Introduction for the General Reader.* New York, 1918

Mill, J. S. *Principles of Political Economy*, various editions

Seager, H. R. *Principles of Economics.* Third edition. New York, 1923

Carver, T. N. *Principles of National Economy.* Boston, 1921

Fetter, F. A. *Economics*, 2 volumes. New York, 1916

Bullock, C. J. *Introduction to the Study of Economics.* New York, 1900

Johnson, A. S. *Introduction to Economics.* Revised. Boston, 1922

LeRossignol, J. E. *Economics for Everyman.* New York, 1923

Bye, R. T. *Principles of Economics.* New York, 1924

Brown, H. G. *Economic Science and the Common Welfare.* Columbia, Mo., 1925

Magee, J. D. *Introduction to Economic Problems.* New York, 1922

Williamson, T. R. *Introduction to Economics.* Boston, 1923

Fairchild, F. R. *Essentials of Economics.* New York, 1923

Marshall, L. C., and Lyon, L. S. *Our Economic Organization.* New York, 1921

II. Selected Readings in Economics

Hamilton, W. H. (Editor) *Current Economic Problems: a Series of Readings in the Control of Industrial Development.* Revised edition. Chicago, 1919

Williamson, T. R. (Editor) *Readings in Economics.* Boston, 1923

Bullock, C. J. (Editor) *Selected Readings in Economics.* Boston, 1907

III. Human Wants

Edie, L. D. *Principles of the New Economics.* New York, 1922

Dickinson, Z. C. *Economic Motives.* Cambridge, 1922

IV. Fundamental Concepts

Fisher, I. *The Nature of Capital and Income.* New York, 1906

V. Statistics and General Statistical Sources

Bowley, A. L. *Elements of Statistics.* London, 1920

Brinton, W. C. *Graphic Methods for Presenting Facts.* New York, 1914

Chaddock, R. E. *Principles and Methods of Statistics.* Boston, 1925

Crum, W. L. and Patton, A. C. *Introduction to the Methods of Economic Statistics.* Chicago, 1925

Fisher, I. *The Making of Index Numbers.* Boston, 1922

Haskell, A. C. *Graphic Charts in Business.* New York, 1922

Mills, F. C. *Statistical Methods Applied to Economics and Business.* New York, 1925

Mitchell, W. C. *The Making and Using of Index Numbers.* Bulletin of the United States Bureau of Labor Statistics, No. 284. October, 1921

Persons, W. M. "Indices of Business Conditions." *Review of Economic Statistics,* preliminary Vol. I, 1919

Young, B. F. *Statistics as Applied in Business.* New York, 1925

Koren, J. (Editor) *The History of Statistics.* New York, 1918

Schmeckebier, L. F. *The Statistical Work of the National Government.* Baltimore, 1925

——— *The Story of the Census, 1790–1916.* Bureau of the Census

Statesman's Year Book, London

Statistical Abstract of the United States, Bureau of Foreign and Domestic Commerce
Abstract of the Census of the United States, Bureau of the Census
Commerce Yearbook, Bureau of Foreign and Domestic Commerce
Yearbook of the Department of Agriculture of the United States
World Almanac, New York

VI. ECONOMIC JOURNALS

The American Economic Review, Cambridge, Mass.
The Quarterly Journal of Economics, Cambridge, Mass.
The Journal of Political Economy, Chicago
The Economic Journal, London
The Annals of the American Academy of Political and Social Science, Philadelphia
The Political Science Quarterly, New York
The Journal of the American Statistical Association, Concord, N. H.
The Journal of the Royal Statistical Society, London
The Economist, London
The Commercial and Financial Chronicle, New York
The Review of Economic Statistics, Cambridge, Mass.
The Federal Reserve Bulletin, Washington

VII. FUNDAMENTALS OF THE MODERN ECONOMIC ORDER

SCHMOLLER, G. *The Mercantile System.* New York, 1896
ROBERTSON, D. H. *The Control of Industry.* New York, 1923
HOBSON, J. A. *The Evolution of Modern Capitalism.* London, 1917
HADLEY, A. T. *The Conflict between Liberty and Equality.* Boston, 1925
MILL, J. S. *On Liberty.* London, 1859

VIII. EUROPEAN ECONOMIC HISTORY

ASHLEY, W. J. *The Economic Organisation of England.* London, 1914
——— *An Introduction to English Economic History and Theory*, 2 volumes. New York, 1910
TOYNBEE, A. *Lectures on Industrial Revolution.* London, 1884
BÜCHER, K. *Industrial Evolution.* New York, 1893
KNOWLES, L. C. A. *The Industrial and Commercial Revolutions in Great Britain during the Nineteenth Century.* London, 1921
DAY, C. *History of Commerce.* New York, 1922
OGG, F. A. *Economic Development of Modern Europe.* New York, 1917
USHER, A. P. *Industrial History of England.* Boston, 1920
GRAS, N. S. B. *An Introduction to English Economic History.* New York, 1922
SLATER, G. *The Making of Modern England.* Boston, 1913
CHEYNEY, E. P. *An Introduction to the Industrial and Social History of England.* New York, 1901

CUNNINGHAM, W. *Growth of English Industry and Commerce*, 2 volumes. London, 1882

SLATER, G. *The English Peasantry and the Enclosure of Common Fields.* London, 1907

SALZMAN, L. F. *English Industries of the Middle Ages.* London and New York, 1923

HEATON, H. *The Yorkshire Woollen and Worsted Industries.* Oxford, 1920

HUTCHINS, B. L. and HARRISON, A. *A History of Factory Legislation.* Westminster, 1911

WEBB, S. and B. *The History of Trade Unionism.* New York, 1920

IX. AMERICAN ECONOMIC HISTORY

JENNINGS, W. W. *A History of Economic Progress in the United States.* New York, 1926

LIPPINCOTT, I. *Economic Development of the United States.* New York, 1921

VAN METRE, T. W. *Economic History of the United States.* New York, 1921

CALLENDER, G. S. *Selections from the Economic History of the United States.* Boston, 1909

BOGART, E. L. *Economic History of the United States.* New York, 1907

COMAN, K. *Industrial History of the United States.* New York, 1905

CLARKE, V. S. *A History of Manufactures, 1607–1860.* Washington, 1916

X. CORPORATIONS AND INVESTMENT

DEWING, A. S. *Financial Policy of Corporations.* New York, 1920

——— *Corporate Promotions and Reorganisations.* New York, 1914

LINCOLN, E. E. *Applied Business Finance.* Chicago, 1924

LOUGH, W. H. *Business Finance.* New York, 1917

GERSTENBERG, C. W. *Financial Organization and Management of Business.* New York, 1924

CONYNGTON, T. *Corporate Organization and Management.* New York, 1917

——— *The Modern Corporation.* New York, 1910

VEBLEN, T. *The Theory of Business Enterprise.* New York, 1904

BROOKINGS, R. S. *Industrial Ownership, its Economic and Social Significance.* New York, 1925

LYON, H. *Investment.* Boston, 1926

JORDAN, D. F. *On Investments.* New York, 1921

LAGERQUIST, W. E. *Investment Analysis.* New York, 1924

KIRSHMAN, J. E. *Principles of Investment.* Chicago, 1924

SAKOLSKI, A. M. *Elements of Bond Investment.* New York, 1921

SMITH, E. L. *Common Stocks as Long Term Investments.* New York, 1925

XI. Marketing

Copeland, M. T. *Principles of Merchandising.* Chicago, 1925
Clark, F. E. *Principles of Marketing.* New York, 1922
Cherington, P. T. *The Elements of Marketing.* New York, 1921
Boyle, J. E. *Marketing of Agricultural Products.* New York, 1925
Weld, L. D. H. *The Marketing of Farm Products.* Minneapolis, 1915
Copeland, M. T. *The Cotton Manufacturing Industry of the United States.* Cambridge, 1912
Cherington, P. T. *The Wool Industry.* Chicago, 1916
Hubbard, W. H. *Cotton and the Cotton Market.* New York, 1923
Buck, N. S. *The Development of the Organisation of Anglo-American Trade, 1800–1850.* New Haven, 1925

XII. Large Scale Production, Combination, and Transportation

See readings suggested at end of Part IV.

XIII. The Economic Functions of Government

See readings suggested at end of Part VI.

PART II

THE LAWS OF PRICE

CHAPTER XI

THE SIGNIFICANCE OF PRICE. DEMAND

In harmony with the title and the general purpose of Part I of this book, the treatment thus far has been descriptive and historical rather than analytical and critical. We have not stopped always to go to the bottom in search of causes and explanations, nor have we gone far into technical details. We have not attempted to answer all of the many important questions that have presented themselves from time to time or the many other equally important questions that have doubtless occurred to the reader. Our purpose has been to lay a foundation for the study of economics in a broad knowledge of the economic organization and its development. From now on our study will be more intensive, more analytical, more critical. We shall have to go more deeply into technical details; we shall undertake to discover the great principles or laws which explain what goes on in the economic world; and we shall seek answers to a multitude of searching practical questions which men are asking each other about our economic organization and its operation.

The significance of price. Attention has already been called to the strategic position which price occupies in the present economic system, and we were at some pains to investigate the influence of price upon production. In so doing however we have by no means attained to a complete picture of the significance of price, for price does far more than merely control production.

Price controls the material enjoyments of life. Everyone is subject to the despotic rule of price, since price controls the kind and the amount of our purchases. And since practically everything that we have comes to us through purchases, price controls our enjoyment of the material things of life. The rich man has many enjoyments because he can pay the prices demanded by those who furnish commodities and services. The poor man has less to enjoy, because he cannot pay the prices. If prices are universally high, everybody, rich and poor, finds his luxuries and comforts, if not his necessities, curtailed. When prices are especially low, everybody feels the pressure relieved and finds himself able to enjoy things which he could not afford before. Individual prices, in their relation to each other, exercise a constant control over the things which we obtain to satisfy our wants. When railroad fares are high, people are compelled to cut down on their pleasure travel. On the other hand, when coal prices soar, we do not cut down much on our consumption of coal; we are forced to make the saving somewhere else, on some article whose use we can better afford to give up or to reduce. A change in the price of a necessity affects the consumption of comforts and luxuries more than the consumption of the necessity itself. Among comforts and luxuries the choice is controlled by price. Polo is the sport of the few, baseball, of the many; not because baseball is inherently a more popular game than polo, but because of the difference in the prices of the respective equipments. When the genius of Henry Ford brought down the price of his product, the automobile ceased to be the rare luxury of the rich and became the common possession of those of moderate means. A generation ago attendance upon the theatre was a luxury to be enjoyed almost exclusively by the well-to-do and only occasionally indulged in by the poor. Since the moving picture industry has made it possible to see an interesting performance for as little as ten or fifteen cents, this enjoyment has become the common possession of the poor. Indeed through changes in price the classification of enjoyments into necessities, comforts, and luxuries (a loose classification at best) is subject to constant revision. There was a time when good woollen and cotton textiles were destined solely for the consumption of the rich and well-to-do. Today the prices of such fabrics are so low that they are used by everybody and regarded as among the prime necessities of life.

Even among necessities there is often choice between substitutes, the decision being determined by their relative prices. Fuel for cooking is a household necessity. But whether the fuel shall be wood, coal, oil, artificial gas, or natural gas is a problem to be determined according to the relative prices of these various fuels. There have been occasions when the farmers of the Middle West burned corn for heating their houses and cooking their food, a choice forced upon them by the relative prices of corn and other fuels. Ordinarily these prices forbid the use of corn as fuel.

From the consumer's point of view then price appears as an all-powerful arbiter, deciding for us what clothing we shall wear, what food we shall eat, what sports we shall enjoy, whether we shall travel or stay at home, whether we shall have many pleasures or few, whether we shall live well or poorly.

Price controls trade. Trade itself is controlled by price. Why was it that the commodities which entered into the long distance trade of the medieval merchants were so generally articles of luxury, jewels, silks, fine wines, spices, perfumes, etc.? The answer is that, transportation costs being so heavy, it was not profitable to bring to market from remote places anything that could not be sold at a high price. Cheap and bulky commodities, such as farm produce, could be sold profitably only in the markets close to the place of production. The modern transportation system is so extraordinarily efficient and modern transportation costs are so low that the province of trade has been enormously extended. But there has been no surrender of the controlling power of price. Goods cannot be brought from one place to another place for sale unless the price at the place of sale exceeds the price at the place of origin by an amount at least equal to the cost of transportation. A fall in the price at the point of consumption may stop the trade in an article formerly brought from a distance. A fall in price at the point of origin will extend the radius governing the territory within which that article can be sold to consumers. Rising prices have opposite effects.

Wages and salaries are prices. Wages and salaries are simply prices, the prices of labor. It is impossible to exaggerate the importance of these prices. To the worker the price which his labor commands is the main factor determining his income and measuring the total amount of his material enjoyments. On the other side of the bargain is the employer. To the manufacturer the price of labor is a cost of production, controlling, as we have seen, the kind and quantity of things which he can produce. To the one who employs personal servants, the price of labor measures the kind and quantity of such service which he can enjoy.

Fluctuation in individual prices. Prices are not fixed and permanent but are subject to constant change. The prices of individual goods change independently of each other, and there are upward and downward movements of the average of all prices — what we call "the general price level." For example the price of wheat in England in the fourteenth century was on the average equivalent to about twenty-five cents a bushel in United States money. In the year 1527 the price suddenly more than doubled. Wheat today is worth $1.25 or more. Here are a few other random examples. The price of codfish in 1891 was more than three times what it was in 1860. Pine boards cost twice as much in 1891 as in 1860; white oak boards cost more than three times as much. The price of sugar on the other hand in 1891 was only a little more than half its price in 1860. Salt in 1891 cost just about half what it cost in 1860. The price of milk was the same in 1891 as in 1860. During the next generation prices of most things increased. For example the wholesale price of salt rose from seventy-nine cents a barrel in 1891 to $2.70 in 1921, three and a half times what it was in 1891. Eggs doubled in price from 1891 to 1921. White pine boards cost nearly five times as much in 1921 as in 1891. Granulated sugar on the other hand increased moderately, the wholesale prices being 4.7 cents a pound in 1891 and 6.2 cents in 1921. Some few things declined in price. Coffee for example cost 16.7 cents per pound wholesale in 1891 and 7.2 cents in 1921, a decline of more than half. Here are the quotations of cotton (middling upland) on the New York market on July 3 of certain selected years from 1893 to 1924, showing a remarkable succession of changes, up and down:

Price of Cotton (middling upland) in New York

1893	$0.08	1916	$0.1295	1922	$0.2375
1899	0.0612	1917	0.2565	1923	0.2725
1914	0.1325	1920	0.3975	1924	0.2975
1915	0.096	1921	0.12		

The most spectacular price changes occur on the speculative markets, such as the wheat market and above all the stock market. Here there are " booms," when all prices go soaring together, and " panics," when all prices come tumbling down, while individual securities rise and fall independently of the general state of the market. The price of a particular stock may double or treble in a few days; while the price of some other security may go steadily down till it reaches zero.

Changes in the general price level. While individual prices may appear to fluctuate without reference to each other, there is nevertheless a tendency of all prices to move up and down as a group. This is what is meant by references to changes in the general price level. Throughout the world's history there has been a strong tendency for prices to increase. It is estimated that prices today are on the average from five to ten times as high as they were 1,000 years ago. In the United States the price level fluctuated moderately during the generation before the Civil War. From 1860 to 1865, prices as a whole more than doubled. Then there followed a steady decline till about 1897. During the next generation, there was an almost uninterrupted rise, reaching its climax in 1920, when prices on the average were more than three times as high as in 1897. This climax was followed by a sudden drop in 1921, and since then prices have stood about twice as high as in 1897. Put in other words, this means that a dollar today will buy just about half as much as it would a generation ago.

The remarkable fluctuations in prices, both of particular commodities and of commodities in general, are a matter of the utmost importance to everyone. To the consumer they affect materially the quantity of goods which his income will buy for him. The business man must constantly keep his eye on price changes. Scarcely any other phenomenon in all the field of economics has so profound a bearing on human welfare.

Importance of future prices. What prices will be in the future is a matter of concern to all of us. When one has to decide whether he can afford a certain expenditure, his decision depends not only on the present price of the thing he contemplates buying, but on his future income (largely a matter of price) and upon the future "cost of living." He has to make the best estimate he can of what these future prices will be. The manufacturer, the merchant, indeed practically every business man, is constantly faced with the necessity of estimating future prices, since the whole fate of any business enterprise may rest upon the course of future prices. The speculators represent a class whose entire activities are concerned with judging future prices and buying and selling goods accordingly, and such speculative dealing comprises an important part of the business on the modern exchanges. There have grown up important business organizations whose function is to furnish their clients with information about past and present prices and prophecies as to future prices. Among such organizations are Babson's Statistical Organization, the Brookmire Economic Service, the Harvard Economic Service, etc. Many of the large city banks furnish this sort of service to their customers.

Complaints about prices. The importance which people justly attach to price is borne out by the frequent complaints which we hear about prices. There is protest about the high prices of certain necessities of life, particularly about the price of meat and other foodstuffs, of coal, and of house rents. From time to time there is complaint that all prices are too high, that the "high cost of living" is getting beyond all endurance, and people hark back to "the good old days" when butter was ten cents a pound, eggs fifteen cents a dozen, and coal five dollars a ton. Such complaints of the high cost of living have been quite the regular thing since the price level began to soar at the end of the nineteenth century. But they are nothing new. There have been popular complaints against the high prices of the necessities of life ever since the Middle Ages, in fact ever since price began to be a matter of importance in the daily life of the common people.

Business men on the other hand are inclined to like high prices, which appear to make business active and profitable. They are more apt to complain of low prices, that "the bottom has dropped out of the market" and it is impossible to make profits. Or the business man may complain that, while prices in general are high, the price of his particular product is unreasonably low. The farmers of the United States complained thus in 1923 and 1924 of the low prices of wheat and other farm products, in the face of a high level of prices in general. Every business man is more or less inclined to see himself thus a special victim of price.

Wages, the price of labor, appear to many to be unduly low. On the other hand there are the big men in industry, railroad presidents, bank presidents, managers of great industrial enterprises, etc., whose salaries (the price of labor again), ranging to heights of $50,000, $75,000, or even more per year, sometimes appear to be too high, out of all proportion to the value of the services rendered. How do these great differences come to be? Is there something wrong? Should not something be done about it?

Should something be done about it? Such questions as the above are frequently arising in connection with complaints about prices. They are not limited to wages and salaries. The farmers of the United States have for generations been inclined to think that something was wrong with agricultural prices, and they have demanded that Congress or the state legislatures do something to relieve them. An understanding of the political history of the United States since the Civil War is impossible without full recognition of this attitude of the middle western farmers toward agricultural prices.

Producers in general, like the farmers, want something done to keep prices high. The general interest of consumers on the other hand is that prices be kept low. So we have demands for legislation or official action to restrict the price of coal, to bring down the price of gasoline, to reduce railway rates, trolley fares, and the rates charged by the local gas company.

Is price a mystery? To some it may appear that price is a mysterious thing, quite beyond human understanding. Some prices appear to be entirely freakish. A diamond may sell for

$1,000, a loaf of bread for ten cents. How is it that the mere luxury is worth ten thousand times as much as a prime necessity of life? A pair of silk stockings may cost five dollars or more. A cotton pair that is stronger and warmer may sell for fifty cents. Why does the price of the less necessary thing exceed that of the more necessary? Such facts as these appear to some to be unexplainable; to others they appear wicked. Is there after all an explanation and a justification?

How is it that prices come to be what they are? We have seen the extraordinary variety of prices. We have seen how they are constantly changing, some rising while others are falling, and again the average of all prices moving steadily downward for years and then as steadily rising. We have seen prices bringing prosperity to some and distress to others. Is all this complexity a mystery incapable of explanation? Is price merely a matter of chance? Or is it the natural result of causes which may be discovered? Certainly price is of enough importance in human affairs to make the search for its explanation worth while if there is promise that the search will be successful.

Prices fixed by statute law or custom. It will appear at the outset that in general prices are not determined by the dictates of statute law. There was a time when prices were largely determined by custom and law. People in the Middle Ages talked of "just prices," "fair prices," etc. In general a price was just if it was as it always had been; the just price was the customary price. When a seller sought to exact a price higher than the customary price he at once laid himself open to suspicion. So also of the buyer who sought to compel a price lower than the customary one. A price different from that established by custom was *prima facie* unjust. Custom was very powerful in those days, and the mere force of custom alone had great influence in preventing serious fluctuations in prices. To the aid of custom came law, civil and religious. Throughout the Middle Ages the law undertook, more or less completely, to prescribe what were the fair and just customary prices and to forbid departure therefrom.

But neither law nor custom was able permanently to prevent the changes in prices which followed in the wake of changed economic conditions. There was some attempt to adapt the law to these changes, and the legal prices were altered from time to time. But the law was unable either to resist the changes or to adapt itself to them. Prices tended more and more to depart from the customary and legal schedules, some higher, some lower. Evidently there were forces at work more powerful than custom and human law.

Legal price-fixing today. The legal determination of price gradually declined and finally ceased to exert an important influence, though the policy has never completely disappeared. Some prices have always been subject to legal determination, even down to the present day. And movements to extend the sway of the law have appeared from time to time in the history of all countries. During the World War, Congress provided for the control of the prices of a great number of important commodities as a war measure. In the United States today legislative authority regulates the rates charged by the railroads for transporting both passengers and freight. Trolley fares, telephone rates, telegraph rates, etc., are subject to legal control. Rates charged by gas companies, electric companies, water companies, and other local public utilities are generally determined or at least limited by legal enactments. Taxicab and hack fares are frequently so limited in the cities. The New York Legislature has recently undertaken to limit increases in house rents. There are a few other examples. But the mention of these cases is sufficient to remind us that they are the exceptions and not the rule. Prices in general are evidently not determined by legislative enactment.

Customary price today. Custom likewise still plays a considerable part in the control of price. The fees charged by physicians, lawyers, and other professional men are quite generally restricted by custom. Custom has an influence on the prices charged by barbers, bootblacks, cab drivers, etc. Many other prices are more or less affected by the force of custom. But it is evident that, in spite of custom, the prices of all these services do change and that there is at no time universal conformity to the customary schedule. What custom does in these cases is not to determine prices, but rather to act as a steadying force, keeping most prices in line with the customary standard and preventing many great and sudden changes. And after all these cases again are the exceptions, not the rule. Prices in general are little restrained by custom today.

Neither sellers nor buyers are generally free to fix prices. We sometimes hear people talk as though they thought the retail sellers of goods had complete power over prices. Does not the merchant, we are asked, fix his prices as he sees fit? What does the customer on the other side of the counter have to say about it? He can take the goods or leave them; that is all. But obviously things are not so simple as this for the retail merchant. In the first place he was a buyer before he became a seller and the prices he charges are very strictly controlled by the prices he had to pay for his goods. Finally the buyer does have an influence on prices, exercised through his power to refuse to buy. The seller cannot put his prices too high; otherwise all or part of the goods will be left on his hands. At the same time it is equally obvious that the buyers do not have entire control of prices. We should not hear so much complaint about prices from both buyers and sellers if either party had the power to determine what prices should be.

Price not a matter of chance. Are prices merely the result of chance? If so, here would be the point for us to stop our study of the subject except for the sake of a historical record of past prices. But the student who has travelled thus far will be pretty sure to answer that question in the negative. There have already appeared too many hints of causal relations here and there. And this answer is right. Prices are not the mere playthings of chance. They are results of the action of great forces or economic laws, laws which may be discovered and understood. The investigation of these laws is the task next before us.

A bargain sale. The first steps in the exploration of the causes of price will prove to be comparatively simple, since they rest upon certain facts of common knowledge. For example a dry goods merchant sees the end of the season for the sale of spring and summer goods approaching, while he still has a large stock of unsold goods on hand. He cannot profitably hold such things over to the next season. They must be sold quickly, if that is at all possible. What does he do? Everybody knows the answer: he cuts prices; a spring "clearance sale" is announced; "bargains" are advertised; goods are displayed in the show windows, with prices plainly marked, something which was not done early in the season. Often the price tag contains the original price, conspicuously crossed out, above or below the new bargain price. Display advertisements in the newspapers, handbills, show window exhibits, etc., are availed of. Everything is done to make the buying public aware of the cut in prices.

Price affects the amount of sales. Of course the explanation of this behavior of the merchant is perfectly simple. Everybody knows that a reduction in price stimulates buying. We know that people generally will choose to buy a larger quantity if the price is low than if it is high, and that many people cannot afford to buy certain things at all if the prices are high, whereas they would buy some at lower prices. A decrease in the price of anything thus generally induces purchasers to buy larger quantities than they otherwise would and also brings into the market purchasers who could otherwise not buy at all.

The converse of this proposition is of course equally true: an increase in prices tends to reduce the quantity that buyers will choose to take. People will generally not buy so much at high prices as at low prices, and certain people, who could afford to buy at low prices, cannot buy at all if prices are high. These generalizations are part of the mental equipment of every merchant.

The demand schedule. This relation between the price of anything and the quantity of that thing which buyers would choose to buy may be illustrated by a schedule, which will show for each price, within the range of possibility, the quantity of the good which people would presumably choose to buy if that were actually the price. For example let the following be the schedule for strawberries in a certain city on a given day.

DEMAND FOR STRAWBERRIES. RETAIL MARKET OF A. MAY 1, 1924

Prices	*Quantities that buyers would choose to take*
55 cents	95 quarts
50 "	100 "
45 "	115 "
40 "	150 "
35 "	200 "
30 "	300 "
25 "	500 "

Market defined. The use of the term " market " in the heading of this schedule makes it pertinent to observe that this term, in the vocabulary of economics, has a somewhat broader content than is given it in ordinary historical descriptions, such as we have followed in the preceding chapters. As technically defined, *a market for any good* (*wealth or service*) *is a place where buyers and sellers exchange that good.* This includes obviously all those markets which we have been studying. It includes also every place where buyers and sellers exchange goods even though there may be no formal market organization. Thus the retail strawberry market of a city need not be any formally organized market in a special building or place; it may be simply the various retail stores and offices where dealers and consumers buy and sell strawberries. It is not even necessary that all the buyers and sellers be physically present at the market. They may be in connection with the market by telephone, telegraph, or mail and represented there by agents or brokers. The market is simply the place at which buyers and sellers make their influence felt, at which exchanges are made, the meeting point of the forces that determine price.

The demand schedule is hypothetical. Let us now see precisely what such a schedule as this means. The number appearing opposite each price indicates the quantity that buyers would choose to take during this particular day if that were the price on this day. It does not show what the price actually is nor what quantity actually is bought. It is in other words a hypothetical schedule. It is based on the presumption of those who are familiar with the market conditions and who are thus able to state with some degree of accuracy about what quantity would be wanted in case any one of these prices should prevail. Such a schedule is ordinarily limited to a few prices above and below the price which is expected to prevail. Theoretically of course the schedule might be continued downward to a price of zero and upward to infinity. Such continuation would be of little value, for two reasons. In the first place we are not able to say with any assurance what quantities would be wanted at the extremes; even the presumptions of the market experts would be little better than wild guesses. And secondly the extremes would serve no useful purpose. All we are interested in is that part of the schedule which relates to prices which might actually be quoted; that is, prices within the range of possibility.

Time and place of the demand schedule. Note also that this schedule relates to a particular time or period of time — to the day of May 1, 1924, or rather to those hours of that day during which buying can take place. This schedule holds true for this day only. The next day will have its own schedule, which may or may not be the same as this one. The prices are alternative prices; only one of them can turn out to be the actual price. So of the quantities; only one can prove to be the amount actually bought. The schedule does not mean that if today 100 quarts are bought at fifty cents, a price of forty cents tomorrow would lead to the purchase of 150 quarts. That would be true only in case tomorrow's schedule should be the same as today's. It may or it may not be.

In the strictest sense the demand schedule does not relate to a *period* of time at all, but merely to an *instant* of time. Theoretically conditions of demand may be changing continuously, so that each succeeding instant has its own demand schedule. For most practical purposes however it is sufficiently accurate to consider demand with reference to a considerable period, during which a certain price holds and a certain quantity is bought. This period, be it remembered, may be quite short, often less than a day or even less than an hour. In the market for certain corporation stocks or certain commodities, such as wheat, cotton, etc., where trading is very active, conditions of demand change rapidly, and there is close approach to the theoretical condition of the constantly changing demand schedule. The demand schedule then relates to a particular time, whether a mere instant or a considerable period, and is not a record of a historical sequence of prices and quantities bought.

Of course such a schedule must relate also to a particular place. Market conditions vary, and a schedule showing the quantities which would be bought at various prices in one market will not hold of any other market, except as a coincidence.

Purpose of demand schedule. We have therefore a schedule showing the numerical relation which is presumed to exist, in a given market at a given time, between the price of a certain good and the quantity that buyers would choose to take. Note again that the numerical statement is a matter of presumption. The general relation between price and quantity wanted — *i.e.*, that the quantity that buyers would choose to take increases with a decrease of price and *vice versa* — is a truth beyond question. It is only when we undertake to express this in exact numbers that we must deal in presumption. In actual business it is seldom if ever possible thus to give exact values to the demand schedule, although the business man is perfectly certain of the general relation between price and the quantity that buyers would choose to take. The problem before us now is to discover the economic laws of price, and for that purpose the construction of such hypothetical demand schedules will prove a useful aid.

Graphical representation. The demand curve. Study of the facts presented in a schedule is generally made easier if the schedule is translated into a curve or graph. Such graphical aids to the presentation and interpretation of statistical material have proved themselves most useful both in the study of economics and in the conduct of practical business affairs. For this reason and for the further reason that we shall have frequent occasion to use the graphical method of presentation in this book, we may well take this opportunity to inquire into the essential features of the most commonly used type of graph.

Referring to the demand schedule above, we note that it presents two separate magnitudes. Our problem is to construct a graph which will disclose the relations between these magnitudes. We must first provide a chart, prepared in the following manner (see Figure 1). Draw a horizontal line parallel to the bottom of the page: this is the *axis of abscissas.* At right angles to this line, through a point on it chosen as the zero position, draw another line perpendicular to the first: this is the *axis of ordinates.* The point of intersection of the two lines is called the *origin;* it is the zero position on each scale. It is customary to letter the origin *o*, and the two axes *x* and *y*, respectively, in order to facilitate reference to the chart. Each axis is divided into a number of equal parts called *scale divisions.* The scale divisions on the two axes need not, and usually will not, be equal to each other. The general considerations determining the choice of scales are first that the entire figure when drawn shall nearly cover the page and second that the inclined lines in the figure shall in general be neither very steep nor too near to the horizontal.

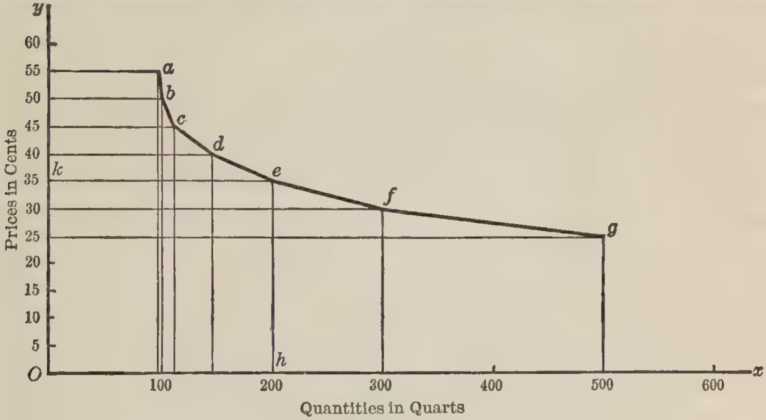


FIG. 1. — Demand for Strawberries. Market of A. May 1, 1924.

The illustration above (Figure 1) shows the two axes prepared in a manner suitable for plotting a curve from the data given above regarding the demand for strawberries in the Market of A. The axis of abscissas (horizontal line, *ox*) is so graduated as to measure numbers of quarts by distances from the origin (*o*) to the right. The axis of ordinates (vertical line, *oy*) measures prices in cents by distances upward from the origin.

It will clarify our explanation of the plotting to emphasize at this point the fact that every point on the chart within the field embraced by the two axes bears a definite relationship to each of these axes. The height of such a point above the base of the chart is measured on the vertical scale and therefore indicates a specific price, while the distance which separates this point from the extreme left-hand boundary of the chart is measured upon the horizontal scale and indicates a certain number of quarts. For example consider the point *e* in Figure 1. By drawing the perpendiculars, *he* and *ke*, from the two axes to this point, we discover that the height of the point above the base axis indicates thirty-five cents (*he* is equal to *ok*) and that its distance from the left-hand border of the chart indicates 200 quarts (*ke* is equal to *oh*). These two perpendiculars are given distinctive names. The vertical perpendicular, *he*, is called the *ordinate* of the point *e*, its length being measured by comparison with the axis of ordinates, *oy*; the horizontal perpendicular, *ke*, is called the *abscissa* of the point, its length being measured upon the axis of abscissas. In like manner it may be seen that every other point on the chart has a twofold significance, its position with reference to the two axes representing a definite portion of each magnitude. It is therefore by a proper placing of these points that the relations between the two magnitudes may be expressed. This truth may be stated in another way; namely, that for every relationship expressed by a pair of magnitudes, consisting of a price and a quantity, there exists one and only one point on the graph which corresponds to this relationship.

Now let us proceed to locate on the chart points representing the respective relations between the seven prices and the seven quantities in the demand schedule above. Starting with the first line of the schedule, we must locate a point whose ordinate measures a price of 55 cents and whose abscissa measures 95 quarts. This is evidently the point *a*. Similarly the point *b* represents a price of 50 cents and a quantity of 100 quarts, and thus are located the other five points. We now connect these points by a series of lines, giving the " curve," *ag*, which is an example of the *demand curve.*

The various points on this curve, *a*, *b*, *c*, *d*, *e*, *f*, and *g*, show the quantities of strawberries (measured on the base axis, *ox*) which it is presumed buyers would choose to take at the corresponding prices (measured on the vertical axis, *oy*). For example, take the price of 35 cents. A horizontal line from the point marked 35 on the *y* axis intersects the curve at the point *e*; a vertical line from *e* intersects the *x* axis at the point marked 200. This means that at a price of 35 cents, buyers would offer to take 200 quarts; or, conversely, if 200 quarts are to be taken, the price must be 35 cents.

Smoothing the curve. Interpolation. Instead of connecting the points on the graph by straight lines, as in Figure 1, it is usual to connect all the points by a curved line passing through them (see Figure 2). This is called " smoothing the curve " and is done chiefly in order that we may consider points on the curve not represented in the schedule. This schedule, for example, went up the scale of prices by steps of five cents. But suppose it is asked: what quantity would buyers presumably offer to take if the price

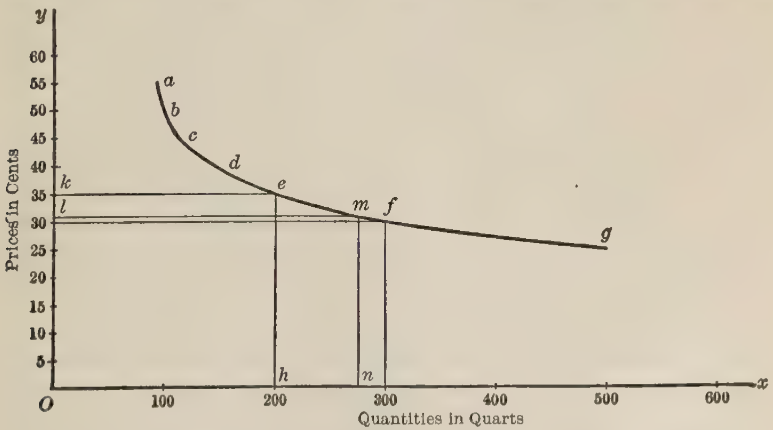


FIG. 2. — Demand for Strawberries. Market of A. May 1, 1924.

were 31 cents? Obviously the quantity would be something between 300 quarts and 200 quarts, the quantities corresponding to 30 cents and 35 cents respectively. This is shown on the curve (Figure 2) by the point m, indicating that at 31 cents the quantity chosen would presumably be 275 quarts. This use of the graph is called *interpolation.*

Definition of demand. We have developed the demand schedule and the demand curve, and the time has now been reached when we should formulate precisely what we mean by *demand.* As a matter of fact, demand is the demand schedule and may be defined thus: *The demand for any good is a schedule of the respective quantities of that good which buyers would choose to take at all possible prices.* The demand curve is simply the graphical representation of demand.

The relation between price and quantity chosen. The quantity of any good which buyers would choose to take is related to, is a function of, the price at which that good may be bought, and it is this relation which is shown by the demand. We must inquire further into the nature of this relation. The demand schedule consists of two columns of figures, one of prices, the other of quantities which it is presumed buyers would choose to take at the respective prices. In the illustrative schedule given above it will be noted that, reading from top to bottom, quantities increase while prices decrease. This means: the higher the price, the less buyers will offer to take; the lower the price, the more buyers will offer to take; or in other words, the quantity that buyers will choose to take varies inversely with the price.[1]

The demand curve brings out clearly this relation between prices and the quantities that would be chosen by buyers, as can be seen from Figures 1 and 2. If prices are measured (as is usual) along the y axis while quantities are measured along the x axis, then, as we go along the curve from right to left, the distances that measure prices grow longer while the distances that measure quantities chosen grow shorter. On the other hand if we travel along the curve from left to right, the distances measuring prices grow shorter while the distances that measure quantities grow longer. This is simply the graphical representation of the principle stated in the paragraph above, according to which the demand curve slopes downward as it goes from left to right.

---

[1] As we shall use the terms, "inversely" and "directly" do not imply that the variation must necessarily be proportioned. If a increases when b increases or decreases when b decreases, we say that a varies directly with b, even though the change in a may not be exactly proportional to the change in b; and correspondingly of an inverse variation.

This is the ordinary relation. But there is one exception. Suppose a state calls for bids to erect a new capitol building. Here there is just one buyer in the market, and there is only one article to be exchanged in any case. The state will take this one capitol building, no more and no less, regardless of the price, up to the limit where the price is so high as to cause the state to give up the building project altogether. What sort of demand schedule and demand curve do we have in this case? In the price column we will list all prices from the lowest that any contractor is likely to bid up to the highest which the state will pay. Opposite each of these figures, in the column of quantities, will of course appear the figure 1. Thus:

Prices	*Quantities that buyers would choose to take*
$6,000,000	1
5,500,000	1
5,000,000	1
4,500,000	1
4,000,000	1
3,500,000	1
3,000,000	1

The corresponding demand curve would evidently be a straight line parallel to the y axis.

The same sort of demand appears when a single buyer calls for bids for a fixed quantity of a certain special article for which there is no demand elsewhere. For example a national government might need 500,000 gas masks of a particular style for its army. It calls for bids. The quantity taken will be exactly 500,000, whatever the price may be, and we have again the vertical demand curve.

The law of demand. Our investigation of demand thus leads to the conclusion that the quantity of any good that buyers would choose to take depends on the price, except when the quantity is a fixed amount. This conclusion may be stated as the *law of demand*, as follows: *In a given market at a given time, the quantity of any good which buyers would choose to take, when not a fixed quantity, varies inversely with the price.*

Elasticity of purchase. Every household uses a certain amount of salt. It would be very uncomfortable to go without salt or even to cut down appreciably upon the usual amount consumed, yet there would be little to be gained by increasing the amount used, since everybody already has about all he wants. If the price of salt should be unusually high, people would still buy about as much as usual. The amount expended for salt is a very small part of the household budget, and even at a much higher price, the cost would be small and the increase could be made up by a slight saving on some other item of expenditure. On the other hand a very low price of salt would not mean any great increase in the quantity bought, for the simple reason that people were already buying about all they wanted. The law of demand holds good of salt, as of other things; that is, at any given time the quantity that buyers would choose to take varies inversely with the price. But it varies only slightly. Considerable variation in price causes only slight variation in quantity. In such a case it is said that purchase is inelastic. The demand curve has a peculiar shape, as shown in Figure 3. Changes in the length of the lines measuring price (as *oa*, *oc*, etc.) are accompanied of course by opposite changes in the length of the lines measuring quantities (as *ab*, *ce*, etc.), but these changes are small. This gives a steep curve.

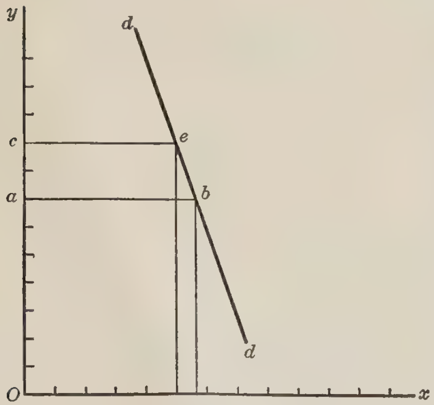


FIG. 3.

On the other hand, there are things which have an elastic purchase. This means that variations in price are accompanied by great variations in the quantities that buyers would offer to take. If the price of a certain high grade automobile were $500 instead of $5,000, the number of persons who would offer to buy that particular make of automobile would be enormously increased. If the price were $15,000, the number of purchasers would be only a small fraction of those who would buy at $5,000. The purchase of any particular kind of car is extremely elastic. The demand curve in such a case has a gentle slope as illustrated in Figure 4. This is because changes in the distances measuring prices (as *oa*, *oc*, etc.) are accompanied by very great changes in the distances measuring quantities (as *ab*, *ce*, etc.). The gently sloping curve is the indication of elastic purchase.

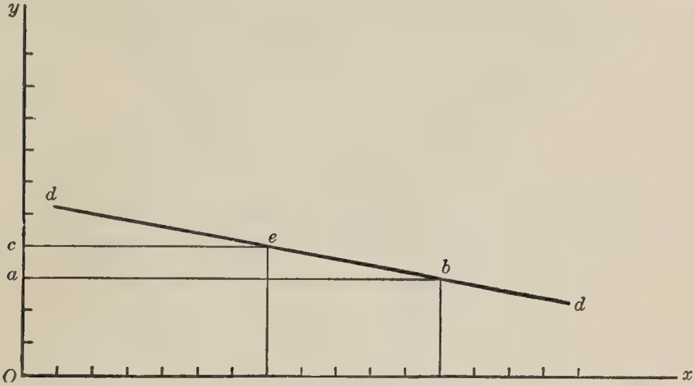


FIG. 4.

Examples of inelastic and elastic purchase. As a general rule inelastic purchase will be found in the case of necessities, since people must have a certain amount, even at the cost of heavy sacrifice of other enjoyments, and since after their needs are fairly satisfied they do not much care for an increased amount. Salt is an example of a necessity. Inelastic purchase is also found in the case of an article for which the desire may be fully satisfied by a small quantity. Salt illustrates this characteristic also. There are other illustrations of goods which are not also necessities. Consider for example the article with which young women improve upon the color imparted by nature to their lips. The quantity of lip sticks purchased depends upon fashion and personal or family opinion; it is not much governed by the price. Few who are inclined to use this beautifier would be deterred by a price even considerably above the prevailing one. And few would be led to more copious use of this aid to nature even though the price should become extraordinarily low. Of course the reason is that anyone can completely satisfy her need for this article with an expenditure of only a small part of her income, no matter what the price may be. The purchase is thus inelastic.

The purchase is generally elastic for luxuries, things that people could do without, especially if the article is also expensive. The theatre is a luxury. The number of tickets that can be sold varies greatly with the price of admission. Witness the immense increase of attendance upon the theatres since the " movies " have made it possible to offer entertainments at prices far below what was possible for the old style show. The purchase of theatre tickets is decidedly elastic. The purchase of washing machines is also elastic. The machine is expensive. Many a family cannot afford one. But the utility of the washing machine is so great and so well recognized that a material reduction in price would certainly lead to a great increase in the numbers purchased. On the other hand an increase in price would soon put the machines beyond the means of many who now purchase them.

" Necessities " and " luxuries." The terms " necessity " and " luxury " are dangerous ones to use in this connection, without careful definition. If by necessity we mean a thing which is required to sustain life or to maintain a decent living in accordance with prevailing standards, then the purchase of necessities will generally be inelastic. But it must be recognized that the purchase is just as inelastic for certain things which are not at all necessities under this definition. And this does not refer merely to such trivial luxuries as the lip stick. Take for example tobacco, which certainly is not a necessity, as we have defined the term. Yet few articles of common use show so inelastic a purchase. The ordinary man will insist upon having his smoke, no matter what the price. If necessary he will make almost any sacrifice in the other items of his personal or family expenditure. The same was generally true of liquors in the United States in the days before prohibition and is true in other countries today.

If our definition of a luxury corresponds to that laid down above for a necessity, then the purchase of luxuries will as a general rule be elastic. But there will be many and important exceptions, of which we have already noted some examples.

Indeed the content of the terms necessity and luxury, no matter how defined, is constantly changing with changes in fashion and in incomes. These terms must therefore always be used with caution, and we must beware of sweeping generalizations dealing with them.

EXERCISES

Note: For convenience in constructing all kinds of charts and graphs, the student is advised to obtain a pad of paper specially ruled in small squares.

1. Draw up an assumed schedule of the demand for coffee on a certain day in a small town.
2. How did you determine the highest price and the lowest price in the schedule of the previous exercise?
3. Construct the demand curve corresponding to the schedule of exercise 1.
4. Take any point on the demand curve of exercise 3, and draw its ordinate and its abscissa. Exactly what does each represent?
5. Take any point on the demand curve of exercise 3 which does not correspond to one of the prices in the schedule. Find the quantity that buyers would presumably take at the price corresponding to this point.
6. State and explain the law of demand. Why does the demand curve normally slope downward as we go from left to right?
7. Draw graphs to illustrate the examples of vertical demand curves mentioned in the text.
8. A city calls for bids to furnish it 200 fire hydrants of a peculiar type nowhere else used. Draw the appropriate demand curve. What determines the two extremities of this curve?
9. Name five articles which you would suppose would have an elastic purchase and explain each.
10. Here are some demand curves (see next page). Which represent elastic purchase and which inelastic?

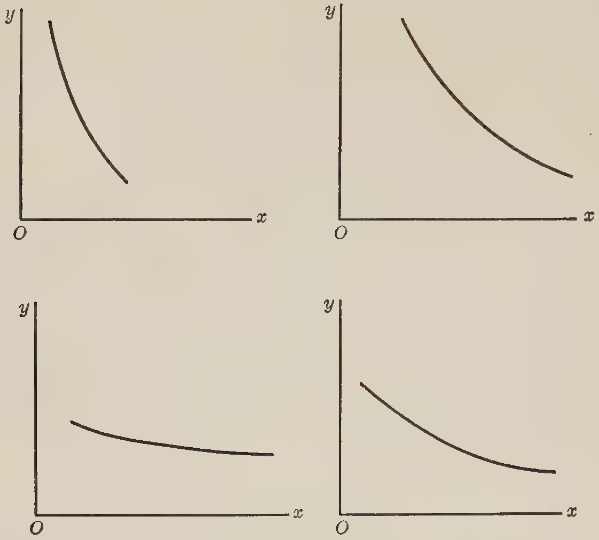


FIG. 5.

CHAPTER XII

SUPPLY

Response of sellers to price. Let us suppose that a stock broker receives an order from a customer to buy a large number of shares of the stock of a certain corporation. The broker goes onto the floor of the exchange and offers to buy at a certain price. Some sellers appear, but the broker soon discovers that sellers are not likely to offer the number of shares which he must buy. What does he do? Of course he offers to pay a higher price. The number of shares offered for sale increases. If it should now appear that he was likely to be offered more shares than he wanted to buy, he would lower his price, thus diminishing the eagerness of the sellers. All this merely illustrates a principle which is common knowledge; *i.e.*, the higher the price bid for anything, the greater is the quantity of that good which will be offered by sellers. Sellers who would part with some of their shares of stock at a low price will sell more if the price is higher. And there are some who would not sell at all at the low price but who can be induced to sell at a higher price.

Supply schedule and curve. The relation between price and the quantity of anything offered for sale may be illustrated by schedules and curves, as in the case of demand. For example the railroads commonly buy cross ties from the farmers in the wooded regions through which their lines pass. The activity of the farmers in cutting trees, shaping the ties, and hauling them to the railroad is directly affected by the price offered by the railroad. Let us suppose that the number of ties that would be delivered to a railroad at a certain point during a particular week at various prices is shown in the following schedule.

SUPPLY OF RAILROAD TIES. FORD'S SIDING. WEEK OF FEB. 18, 1924

Prices	*Quantities that would be offered*
$1.00	9,000
.90	8,500
.80	8,000
.70	7,000
.60	6,000
.50	4,500
.40	3,000
.30	1,000

The corresponding curve is shown in Figure 6 below:

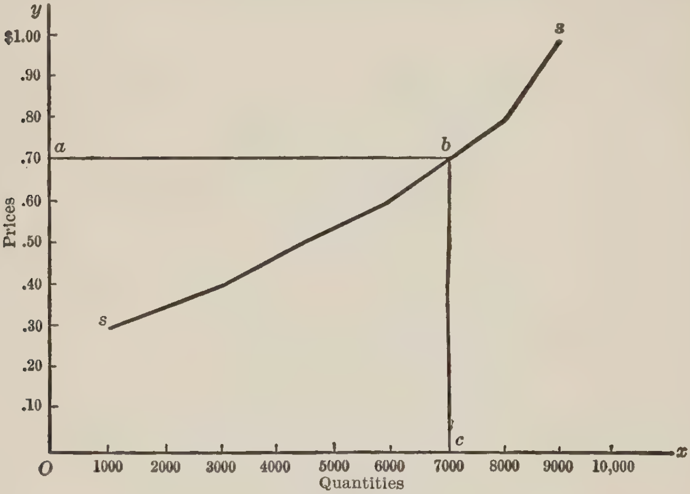


FIG. 6.

Supply relates to a particular time and a particular place. In our study of demand schedules and curves it was seen that any such schedule or curve must always relate to a particular time and place. The same is true of supply. The schedule and the curve which we have just constructed are presumed to represent the conditions of supply of a certain good (railroad ties) at a certain place (the railroad station at Ford's Siding) and at a certain time (the week of February 18, 1924). Similar schedules and curves might be constructed for other goods and for other markets and other dates. The conditions of supply of railroad ties would not necessarily be the same at another point on the railroad even during this same week, nor would they necessarily be the same at Ford's Siding at some other time, as the week before or the week after. In strictest theory a supply schedule or curve (exactly as in the case of demand) relates to an instant of time; the next instant the conditions may be different. For most purposes however the time is an appreciable period, a few minutes, an hour, a day, a week, etc., during which the respective quantities would presumably be offered if various assumed prices should prevail.

Supply schedules and curves are hypothetical. We should remind ourselves here of another respect in which supply corresponds with demand. The schedule and the curve are hypothetical, in that they assume certain quantities to illustrate a relationship whose general character we know without knowing the exact numerical quantities. The market expert may be able to estimate very closely, in certain cases, the quantities that would be offered at various prices. No one can be absolutely certain of these quantities. But we do know the normal relation between price and quantity offered; we know that if, at the actual price, a certain quantity was offered for sale, a higher price would have brought out a larger quantity and a lower price would have brought out a smaller quantity. The railroad company in our illustration can announce a certain price for railroad ties, say seventy cents, and find out readily enough how many the farmers will offer to sell at that price, say 7,000. Having fixed that price for this week, it will never know how many ties would have been offered at some other price, say eighty cents or sixty cents. But it is practically certain that at eighty cents the number offered would have been more than 7,000 and that at sixty cents something less than 7,000 would have been forthcoming. This relation between price and quantity offered is all that the supply schedule is intended to show. The exact numerical quantities assumed, though purely hypothetical, are entirely satisfactory for our purpose; namely, the inquiry into the economic nature of supply.

The definition of supply. The definition is analogous to that of demand. *The supply of any good is a schedule of the respective* *quantities of that good which sellers would offer at all possible prices.* Supply is simply the supply schedule, and the supply curve is the graphical representation of supply.

The relation between price and what sellers would offer. The law of demand was developed in the last chapter, and we may now proceed to derive a similar law for supply. In the supply curve, as in the demand curve, we usually measure prices on the y axis and quantities on the x axis (see Figure 6). As we go along the curve from left to right, the distances that measure prices grow longer and the distances that measure quantities offered also grow longer. This simply means the higher the price, the greater the quantity offered. On the other hand as we travel along the curve from right to left, both the distances measuring prices and those that measure quantities offered grow shorter; which is the graphic expression of the principle, the lower the price, the smaller the quantity offered. This relation is portrayed by the upward slope of the curve as it goes from left to right, and it is because of the principle that the quantity offered for sale generally varies directly with the price that the supply curve thus slopes upward.

This is the ordinary relation. But it sometimes happens that the quantity that sellers have to offer is fixed and not subject to variation in relation to price. A familiar example is an auction sale of a particular article, as a painting by an old master. There is only one such article for sale no matter what the price may be. The supply schedule shows 1 opposite each price, and the supply curve is a straight vertical line, beginning at the lowest price at which the owner would sell and extending to the highest price which any buyer may be supposed to be willing to pay. A similar curve results when a single seller has a fixed quantity of an article which is nowhere else obtainable and which he proposes to sell all in one lot to the highest bidder. After the World War the United States Government disposed of many lots of army supplies in this way.

Lest we overlook any conceivable case, notice should be taken of the possibility of a third form of supply curve, which may occur in connection with the offer of goods or services by extremely improvident sellers. Thus traders dealing with savages have sometimes sought to stimulate the natives to increased activity in bringing in their products by the offer of higher prices, only to find the result disconcertingly contrary to the normal law of supply. The native was content when he had earned a certain amount; thereafter he proposed to knock off and enjoy his leisure. The higher the price, the fewer the goods he had to sell to reach his goal. His supply curve, after ascending normally for a while, turned back on itself (see Figure 7). This situation may conceivably exist in case of a civilized producer, either because he is unusually improvident or because his product or service is unique and very highly paid. But evidently such situation can have no effect on the total supply of any important commodity or service and is of no practical importance.

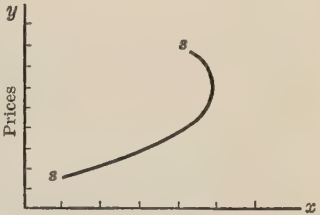


FIG. 7.

The law of supply. The study of supply thus leads to conclusions which may be expressed, with full allowance for such possible exceptions as are illustrated by the backward supply curve, in the *law of supply* as follows: *In a given market at a given time, the quantity of any good which would be offered for sale, when not a fixed quantity, generally varies directly with the price.*

Elasticity of sale. There is variation in the elasticity of the sale of different commodities. This subject is analogous to the elasticity of purchase, which was discussed at some length in the last chapter, and the subject of elasticity of sale may therefore be dismissed with a few words here. When the quantity of a good offered for sale changes greatly with changes in price, its sale is said to be elastic. The corresponding supply curve will rise slowly from left to right; its slope will be gradual, as in Figure 8.

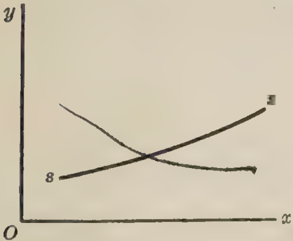


FIG. 8.

Sale is inelastic when price variations are accompanied by only small variations in the quantity offered for sale. The supply curve in such cases has a steep slope, as in Figure 9.

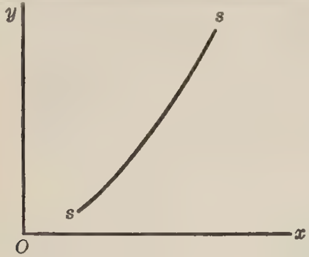


FIG. 9.

In general the sale of the products of the extractive industries is inelastic, whereas manufactured articles have an elastic sale. These rules are by no means without exception, and this whole matter will be more easily studied in connection with our investigation of the costs of production in a later chapter.

Are the laws of demand and supply universal? It may have been noted that the word "generally" was inserted in the statement of the law of supply, though not in the law of demand. This suggests the question of the universality of these laws. The possibility of exceptions comes from the backward action of price upon demand and supply, and the investigation of this problem will be much easier after we have studied the laws of price. For the present it may merely be suggested that there are certain recognized exceptions to the law of supply, of which we have taken notice. These exceptions are of slight practical importance. Their presence is recognized by the word "generally" in the statement of the law, but the law may be assumed to hold in virtually all practical cases. In the case of demand, while it may be possible to imagine fanciful cases of exceptions, there appear to be no real exceptions to the law. This matter will be further examined in a later chapter.

EXERCISES

1. Draw up an assumed schedule of the supply of potatoes in a small city market.
2. How did you determine the highest price and the lowest price in the schedule of the previous exercise?
3. Construct the supply curve corresponding to the schedule of exercise 1.
4. Take any point on the supply curve of exercise 3, and draw its ordinate and its abscissa. Exactly what does each represent?

5. Take any point in the supply schedule of exercise 3 which does not correspond to a price in the schedule. Find the quantity that sellers would presumably offer at the price corresponding to this point.

6. Look up the stock market report in any newspaper, and note the number of shares of the stock of the Pennsylvania Railroad sold on that day, also the opening price, the closing price, the high price, and the low price for the day. Construct what you would consider a probable supply schedule for this stock at the opening of the exchange on that day.

7. State and explain the law of supply. Why does the supply curve normally slope upward from left to right?

8. A bankrupt nobleman decides to sell at auction a unique painting by an old master. Draw the appropriate supply curve. What determines the two extremities of this curve?

9. Name five articles which you would suppose would have an inelastic sale, and explain each.

10. Can you reach any conclusion, from reading the newspaper market reports, as to whether the sale of wheat in a large city market is elastic or inelastic? How about the sale of United States Steel Corporation stock?

[illegible] adequate the
supply [illegible]
The demand.

[illegible]

[illegible]
[illegible] bought by auction [illegible]

[illegible]

CHAPTER XIII

COMPETITIVE PRICES. THE DEFINITION OF DEMAND AND SUPPLY

In the early days of the science of economics a cynical observer is reported to have said: "If you want to make a first class economist, catch a parrot and teach him to say 'supply and demand' in response to every question you ask him." It must be admitted that there is a common impression that the whole solution of the problem of price is contained in the statement that "price is determined by supply and demand." It is true that the prices of most goods, though by no means of all, are thus determined, but the problem of how prices are determined is not quite so simple as this statement would imply or as is commonly assumed.

Competition. At the very outset of our analysis of the causes determining price, we are confronted by the important factor of competition, and we shall find it helpful to divide prices into two groups, (1) competitive prices, (2) non-competitive prices, and to devote our attention at first exclusively to the former group. This requires a definition of competition more precise than that which has sufficed us thus far. *Competition is the effort of two or more sellers to sell to the same person or persons, each seller acting independently in his own interest without regard to the interests of other sellers.*[1]

Competitive price is a price determined under conditions of competition of sellers. Competitive prices are determined (1) by demand and supply or (2) by the decision of a buyer or (3) by cost of production. Our problem is to discover how these various forces operate to determine competitive prices. Incidentally we shall learn how the quantity that is exchanged is at the same time determined. We shall commence with the investigation of the first and most important group; *i.e.*, competitive prices determined by demand and supply.

---

[1] This is, obviously, competition of sellers, and there may be also competition of buyers, with an analogous definition. The latter concept is however seldom employed, and where the term competition is used without qualifying words competition of sellers is to be understood.

A concrete case for illustration. In order to have the assistance of a concrete example, let us assume that we are investigating the price of fresh-laid eggs in the retail market of a small-sized city, which we shall call A. We will assume that on a certain day, October 15, 1924, the demand for eggs and the supply of eggs were as shown in the following schedules:

Retail Egg Market of A. October 15, 1924

DEMAND		SUPPLY	
Prices	*Quantities that would be taken*[1]	*Prices*	*Quantities that would be offered*
$1.00	145 dozen	$1.00	400 dozen
.95	150 dozen	.95	390 dozen
.90	160 dozen	.90	375 dozen
.85	170 dozen	.85	360 dozen
.80	185 dozen	.80	340 dozen
.75	200 dozen	.75	320 dozen
.70	220 dozen	.70	300 dozen
.65	240 dozen	.65	280 dozen
.60	260 dozen	.60	260 dozen
.55	280 dozen	.55	230 dozen
.50	300 dozen	.50	200 dozen
.45	320 dozen	.45	160 dozen
.40	340 dozen	.40	120 dozen
.35	360 dozen	.35	80 dozen
.30	380 dozen	.30	40 dozen
.25	410 dozen	.25	
.20	440 dozen	.20	
.15	480 dozen	.15	
.10	540 dozen	.10	
.05		.05	
.00		.00	

The hypothetical character of such schedules has been made clear in the two previous chapters. The relations between the quantities and the prices conform to the well-known laws of demand and supply. The exact numerical quantities are not significant; they are used only for the purpose of illustrating the laws.

---

[1] Hereafter, when convenient, we shall use this elliptical phrase for the more precise one, "quantities that buyers would choose to take."

Now let us plot on one diagram two curves corresponding respectively to these two schedules. The result is shown in Figure 10; *dd* is the demand curve; *ss*, the supply curve.

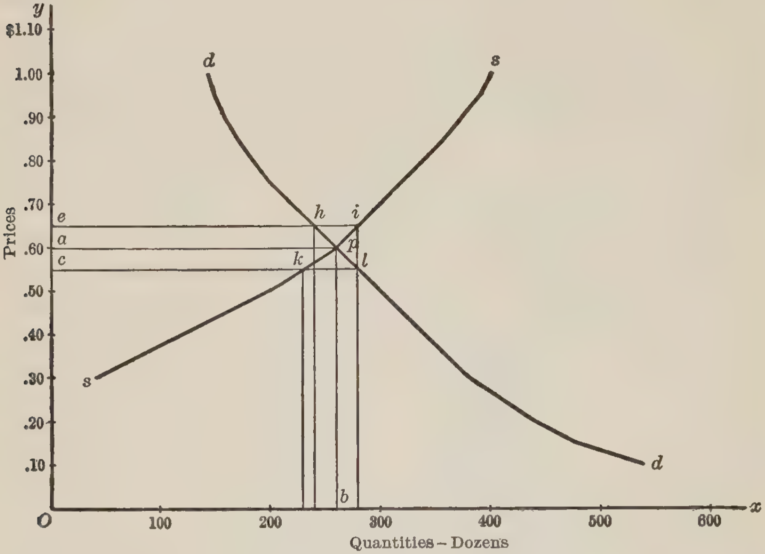


FIG. 10. — Retail Egg Market of A. October 15, 1924.

The market price: quantities offered and taken are equal. The curves of demand and supply intersect at the point *p*. What does this mean? In the first place it is evident that the ordinate of this point, *pb*, represents the same price on both curves. If we are considering the demand curve *dd*, *pb* represents a price of sixty cents. It also represents a price of sixty cents if we regard *p* as a point on the supply curve *ss*. Also the abscissa of the point *p* represents the same quantity on each curve; *i.e.*, 260 dozen. The point *p*, as a point on the curve *dd*, indicates that at a price of sixty cents the quantity of eggs that would be taken by buyers in A is 260 dozen; as a point on the curve *ss*, *p* indicates that at a price of sixty cents the quantity of eggs that would be offered by sellers in A is 260 dozen. At the price of sixty cents the quantities offered and taken would be the same.

In the second place we note that sixty cents is the only price at which these quantities can be equal. For example if the price were sixty-five cents, the quantity offered would be 280 dozen, represented by the line *ei*, whereas the quantity that buyers would take would be 240 dozen, represented by the line *eh*. The quantity offered would be greater than the quantity taken, and this is clearly true of every price greater than sixty cents. On the other hand consider any price lower than sixty cents, as fifty-five cents. At that price the quantity offered is 230 dozen, represented by the line *ck*, while the quantity that would be taken, represented by the line *cl*, is 280 dozen. The quantity offered is now less than the quantity taken, and this will obviously hold of every price lower than sixty cents.

The price at which exchanges take place. This price is thus distinguished from every other possible price. It is the price at which exchanges will take place. And if we assume a real market, in which there are several buyers and sellers, each fairly familiar with the conditions of demand and supply of eggs on this particular day, this price is the only price at which any considerable amount of eggs will be exchanged. A seller who would demand more than this, say sixty-five cents a dozen, would find nobody willing to buy from him. Rather than be left with his stock unsold on his hands, he would reduce his price. A seller on the other hand who ignorantly offered to sell eggs for less than sixty cents, say fifty-five cents, would soon find himself swamped by a rush of eager buyers; he would see that he could easily sell all his stock for more than fifty-five cents and he would raise his price. We reach the same conclusion on regarding the matter from the buyer's standpoint. A purchaser who ignorantly offered to buy any considerable quantity at more than sixty cents a dozen would find all the sellers eager to sell to him and would soon see that he could get all the eggs he wanted at a lower price. He would therefore lower his price. Any buyer on the other hand who tried to get eggs for less than sixty cents would find no one willing to sell to him and would have to raise his price or go without. No considerable buying and selling could therefore take place at any price except sixty cents, so long as our curves accurately represent the conditions of demand and supply.

It is of course quite probable that there will be sales at other prices. It may take some appreciable time for the actual market to settle down to a state of equilibrium, during which time there will probably be scattered sales at various prices. Even after equilibrium is reached occasional ignorant buyers or sellers, such as we have referred to, will very likely make some purchases or sales at unfavorable prices before discovering their mistake. It is often convenient to use the term "market price" to distinguish the price, determined by the intersection of the curves of demand and supply, at which the bulk of exchanges are taking place from the off prices at which some few scattered sales may be made.

The price that clears the market. The market price (sixty cents a dozen in our example) is unique in another respect. At this price all sellers of eggs are able to dispose of the entire quantity of eggs which they are willing to sell (at that price) and all buyers are able to get as many eggs as they wish to buy (at that price). No seller who is willing to take sixty cents a dozen will be left with an unsold stock, and no buyer who is willing to pay sixty cents will fail to get what he wants. In technical terms, this price "clears the market." This is not true of any other price. If, assuming for the moment the impossible, the market price were sixty-five cents, all buyers would obtain all they wanted at that price; *i.e.*, 240 dozen; but sellers, willing to dispose of 280 dozen, could sell only 240 dozen, because buyers would take no more at that price. The difference, forty dozen (represented by *hi*, the difference between *ei* and *eh*), would be left on their hands. On the other hand consider a market price of fifty-five cents, again assuming the impossible. Sellers would be able to dispose of all they had for sale at that price; *i.e.*, 230 dozen. But buyers, although willing to buy 280 dozen, would be able to get only 230 dozen, since the sellers would part with no more. The difference, fifty dozen (represented by *kl*, the difference between *cl* and *ck*), stands for the unsatisfied requirements of the buyers. Such differences cannot exist at the price *pb*, determined by the intersection of the demand and supply curves.

The case of the vertical demand curve. The example which we have used represents the ordinary cases of competitive price, in which both demand and supply curves have the normal shape. No new difficulty arises when we consider cases in which there is the vertical demand curve. For example a government calls for bids for a certain number of rifles based upon its own peculiar specifications which are acceptable to no other user. Assuming that two or more manufacturers desire the contract, we have competition of sellers, but only one buyer. The appropriate demand and supply curves are shown in Figure 11. Their intersection at *p* determines the price *pd′* at which the quantity desired by the government *pa*, will be exchanged. The price is thus determined by demand and supply exactly as in the previous case of the egg market. The important difference between the two cases is that in the first case demand and supply also determined the quantity exchanged, whereas in the second case the quantity was fixed in advance by the demand when the government determined that it would buy just so many at any price from *o* to *d′d*.

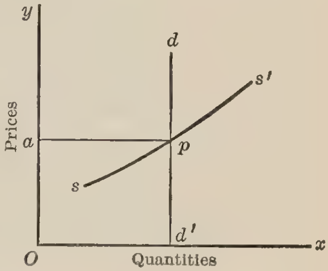


Fig. 11.

Generalizations. From these examples [1] we may pass to certain general conclusions regarding competitive market price when determined by demand and supply. The price thus determined is that price which clears the market; it is that price at which the quantity offered by sellers is equal to the quantity that buyers would choose to take; and this quantity is therefore the quantity that is actually exchanged. Graphically represented, the price is determined by the point at which the demand and supply curves intersect, the price being usually measured by the ordinate of the point of intersection. The quantity exchanged is likewise determined by the intersection of the two curves except when the demand curve is a vertical line, in which case that curve alone determines the quantity.

---

[1] If the question should be raised why we do not first complete our analysis by considering also the vertical supply curve, the answer is that this curve does not represent a situation of competition but pertains to monopoly, a subject which will have our attention in a later chapter.

Price fixed by a sole buyer. This is a convenient place to investigate the second group of competitive prices; namely, that in which the price is determined by the decision of a sole buyer. Suppose a canning establishment is set up in a country district and that the manager advertises that he will pay five cents a quart for all wild berries brought to the factory. We assume also that there is no other local market for wild berries. Here we have a market in which there will be the normal supply curve representing competition of sellers, but no demand curve; *i.e.*, no curve showing for different prices the quantities that would be taken by the buyer. Instead we have a price fixed by the buyer, who is ready to take any quantity at that price, or any quantity up to a certain limit. The result is illustrated by Figure 12. There is the normal supply curve *ss*. The price was fixed in advance at five cents. The quantity exchanged is the result of this price and the supply, determined graphically by drawing a horizontal line from *a* to meet the supply curve at *p*. The quantity exchanged is measured by *ap*. If the buyer set a limit to the quantity he would take, that is measured by the length of the line *ab*.

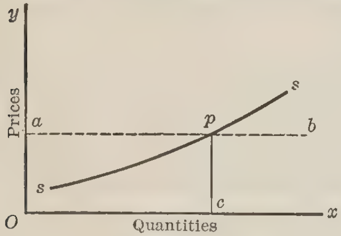


FIG. 12.

Bringing together the two cases last examined (Cf. Figures 11 and 12), we note that a sole buyer can decide either the price he will pay or the quantity he will take; he will not decide both, and, having decided the one, he will let the other be determined by the operation of supply.

Competitive price fixed by cost of production. In certain lines of industry, as will later be developed, the market price is, between certain limits, determined solely by the cost of production. This situation is, as regards our present inquiry, somewhat similar to the preceding case of the price fixed by a sole buyer. We now have a normal demand curve (*dd* in Figure 13), but no supply curve. Instead there is the horizontal line *cc′*, a cost curve, which indicates that any quantity between *oa* and *ob* will be offered at a price of *oe*. This price is therefore determined, and it only remains to note that the quantity exchanged is determined by the intersection of the demand curve and the cost curve. The investigation of the causes which produce this particular cost situation will have to be deferred to a later chapter (Chapter XVII).

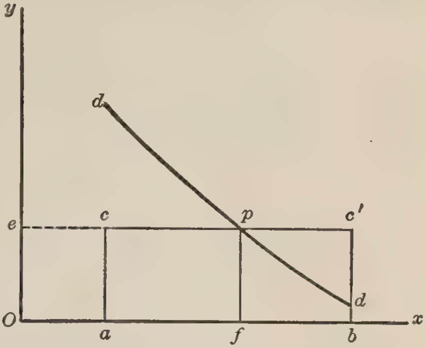


FIG. 13.

Defining demand and supply. Before proceeding further in the investigation of the forces determining price, we shall find it advisable to pause here for a somewhat critical consideration of our definitions of demand and supply and a closer scrutiny of the nature of demand and supply.

Two possible meanings of demand. The whole investigation of the laws of price is in danger of being obscured by the fact that the term "demand" is used with two different meanings in popular speech and writing and, it must be admitted, in much scientific writing also. A clear understanding of these two meanings will enable the student to avoid much confusion in his later reading and discussion. For example we hear "there is a heavy demand for wheat"; "the demand for high shoes is decreasing rapidly"; "the demand for automobiles is unusually strong this spring"; etc. In all such expressions it is evident that the word "demand" refers to the whole *demand schedule* (or curve). The meaning is that at any given price the quantity that buyers would choose to take is large or small, greater or less, etc. In Figure 14, the curve *d′d′* represents a heavy demand as compared with the curve *dd*, since whatever price is chosen the quantity that would be taken according to *d'd'* is greater than according to *dd*. Thus if *dd* represents the demand for coal at a given time, a price of $15 would on this day induce buyers to take 100 tons; if on some other day the demand were as shown by the curve *d'd'*, this means that buyers would on that day take 250 tons at $15 a ton. The same result would follow if any other price were taken and the appropriate ordinates and abscissas drawn. The same notion of demand is likewise in mind when such expressions as these are used: "the price of anything depends on demand and supply"; "the greater the demand, the higher the price"; etc.

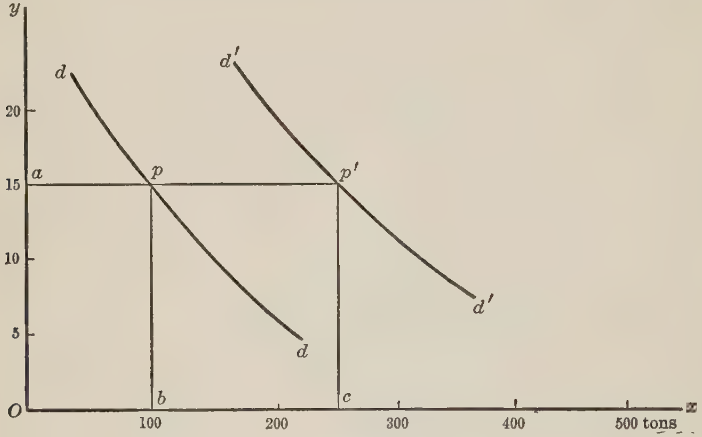


Fig. 14. — Demand for Coal.

The foregoing usage evidently agrees with the definition of demand which we have accepted (see Chapter XI). On the other hand we shall meet such statements as these: "an increase in price causes a decrease in demand"; or "the higher the price, the smaller the demand." Now it is evident that these statements are in direct conflict with the last quotation in the previous paragraph. Something must be wrong when people say "the greater the demand, the higher the price" and, in the next breath, "the higher the price, the smaller the demand." The trouble comes from shifting the meaning of the word "demand." In the second expression demand does not mean the whole demand schedule or curve. It means rather *the quantity that buyers would choose to take at some particular price.* Refer again to Figure 14, and note that, according to the curve *dd*, at a price of fifteen dollars 100 tons of coal would be taken. This quantity, 100 tons, is the demand for coal at fifteen dollars, when the word is used in this second sense. It is represented by the line *ap* or the line *ob*. When we talk of an increase or a decrease in demand in this sense, we do not mean a change in the whole demand schedule or curve. We mean rather a movement along the curve, locating a different point, whose distance from the *y* axis (which measures demand in this sense) is greater or smaller than before. When used in this sense we can never think of demand except in reference to *a particular price,* stated or implied. We cannot speak simply of the demand for cotton. We must say: the demand for cotton at fifteen cents a pound, or at some other price.

Choice of a definition of demand. Now it is idle to argue about definitions; and this presentation of the two senses in which the term demand is used was not offered as an introduction to a battle of words. As has been previously stated (Chapter I), there is no absolute test of the correctness of a definition. Either of the two possible meanings of the term " demand " might be chosen. Both are about equally favored in the loose usage of popular speech and writing. To that extent we violate popular usage when we choose to make exclusive use of either meaning; but that is inevitable. We are free then to select the meaning which will best serve the purpose of the scientific investigation of economic laws. The central problem in economics is concerned with the laws which determine price and value. In the study of this problem, while it would be possible to make fairly satisfactory progress with either definition of demand, the advantage is decidedly in favor of the first meaning. We have therefore accepted as our definition of demand the following: *The demand for any good is a schedule of the respective quantities of that good which buyers would choose to take at all possible prices.* The justification of this choice will appear more clearly as the reader proceeds through these pages.

Two possible meanings of supply. The term “supply” is commonly used in two senses, corresponding exactly to the two meanings which are given to “demand.” When people say “the supply of cotton is short this year,” “the wheat supply is greater than ever before,” etc., they are using the term as the whole *schedule of quantities that would be offered at various prices.* They mean for example that, at whatever price bid, the quantity of cotton which sellers would offer is small, relative to the quantities that would ordinarily be offered. The term is used also in this sense, when it is said: “The price of anything depends on demand and supply”; “the greater the supply, the lower the price”; etc.

This usage is obviously in harmony with the definition of supply which we have adopted (see Chapter XII). However such expressions as these will frequently occur: “An increase in price causes an increase in supply”; “the supply depends on the price”; “the lower the price, the smaller the supply.” There is here the same confusion as results from the two meanings of “demand.” For example the last quotation above is clearly in flat contradiction with the last quotation of the previous paragraph. As clearly, the trouble is due to shifting the meaning of the term “supply.” In the sentences quoted in this paragraph, “supply” does not mean a schedule of quantities that would be offered at various prices; it means *the quantity that would be offered at some particular price.* Referring to Figure 15, it will be seen that, at a

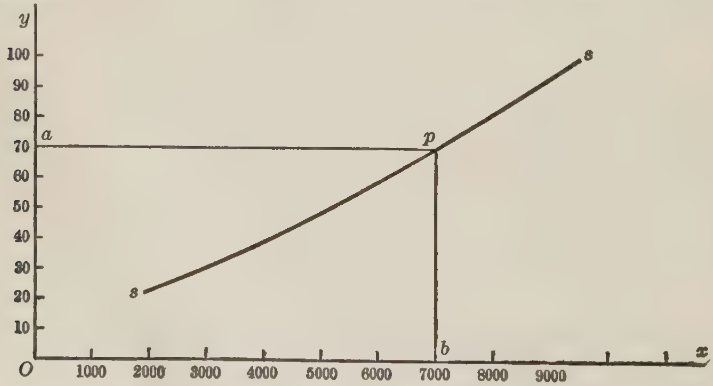


FIG. 15. — Supply of Railroad Ties.

price of seventy cents, 7,000 railroad ties would presumably be offered. In the second sense of the term " supply," it is this quantity that is the supply of railroad ties at seventy cents. It is represented by the line *ap* or the line *ob*. When we speak of an increase or decrease in supply in this sense, we mean, not a change in the whole supply curve, but rather a movement along the curve, thus locating a different point whose distance from the y axis (which measures supply in this sense) is greater or smaller than before. Supply, in this sense, is always in reference to a particular price. We cannot conceive of the supply of wheat in general. It must always be the supply at $1.50 a bushel, or some other price.

Choice of a definition of supply. All of this corresponds evidently to our study of the two meanings of the term " demand." Here also it is true that fairly satisfactory progress might be made with either definition but that the advantage is distinctly with that which regards supply as the whole schedule rather than a particular quantity offered. Hence we have adopted this definition: *The supply of any good is a schedule of the respective quantities of that good which sellers would offer at all possible prices.*[1]

The importance of consistency. While satisfactory results might come from the employment of either of the possible definitions of demand and similarly of supply, nothing but confusion can result from a vague use or from a mingling of the two meanings which might be attached to each of these terms, twin sins of which it must be confessed that some very good economists are occasionally guilty and which run riot in popular speech and writing. Equally essential is it that the definitions chosen for demand and supply be in harmony with each other. That is, both terms must be defined either in the schedule sense or in the sense of the quantity taken or offered at a particular price. We cannot define demand for example as a schedule of quantities that would be bought, while defining supply as the quantity that would be offered at a particular price. We have chosen, for reasons stated and which will become more apparent as our study proceeds, the schedule sense for both terms. These definitions will be adhered to consistently throughout this book. The reader who goes on with his study of economics is sure to meet writers whose usage is different. If he has mastered the lesson here presented, he will have little difficulty with those writers who have chosen and used consistently the alternative definitions. He will have more difficulty in understanding those writers whose usage is vague or who mix the two definitions, but even here his task will be lightened if he has a clear notion of what the two possible meanings are and how they may be used.

---

[1] Our choice of these definitions of demand and supply compels us to adopt different terminology for what is ordinarily known as elasticity of demand or supply. As we use the terms it is not, for example, demand that is elastic, but "the quantity that buyers would choose to take," since it is this that varies in accordance with price. This is the reason for our adoption of the terms "elasticity of purchase" and "elasticity of sale" when treating of this subject in Chapters XI and XII.

Demand and desire. We cannot safely leave this discussion of definitions until we have taken precautions against two pitfalls of confusion which stand ready to bring the unwary to grief. Demand is not the same as desire. An impecunious student, standing with empty pockets outside a theatre, may have a keen desire to see the show, but we cannot say that he has a demand for a ticket. His desire has no effect upon the demand for tickets, no effect upon the numbers that could be sold at various prices. To have a demand for anything one must not only desire it, want it; he must also be able and willing to buy. Most of us desire fine automobiles, trips to Europe, and other things without number. Most of these desires have no effect upon the demand for these things. The only one who contributes to the demand for a certain make of automobile is the one who is able and willing to buy at least one at a price at which this automobile is likely to be for sale. We have already paid some attention to the interesting subject of human desires or wants (Chapter I) and we shall later see how closely they are related to demand. What we need now is to be on our guard against thinking desire and demand are the same thing.

Supply and stock. The second pitfall that threatens is the confusion of supply with the stock of goods on hand. The elevators in the neighborhood of Chicago contain millions of bushels of wheat of various grades. This stock of wheat is not the supply of wheat on the Chicago Board of Trade. The supply is the schedule of the respective numbers of bushels of wheat that would be offered for sale at various prices at a particular time. These quantities will not be the same as the quantity of wheat in the Chicago elevators. In fact, while there is a relation between the quantity of goods on hand and the supply, this relation will often appear not to be very close. The quantity of wheat offered for sale at Chicago at some particular price may be far less than the wheat on hand in the elevators. Much of that wheat may not be for sale at that price. On the other hand the quantity offered might be more than the total stock in the elevators. The Chicago market may be drawing upon wheat stored in other places. In fact the quantity offered might conceivably be greater than the total stock in existence. This surprising state of affairs occurs now and then in speculative markets. On May 9, 1901, on the New York Stock Exchange, it was suddenly discovered that sellers had contracted to deliver more shares of the stock of the Northern Pacific Railway Company than could possibly be bought or borrowed by them. A disastrous crisis was the result. Of course, the people were offering shares which they did not then possess; they hoped to acquire them in time to deliver. But that did not change the fact that so many shares were actually offered.

EXERCISES

1. The demand and supply schedules for sugar in the retail market of a certain city on a certain day are presumed to be as follows:

Prices	*Quantities that would be*	
Cents	*taken*	*offered*
5	900 pounds	360 pounds
6	700 pounds	460 pounds
7	540 pounds	540 pounds
8	440 pounds	600 pounds
9	360 pounds	640 pounds
10	320 pounds	660 pounds

(*a*) Draw the corresponding demand and supply curves.
(*b*) What is the price of sugar?
(*c*) Why can the price not be lower?
(*d*) Why can the price not be higher?
(*e*) Would it be correct to say that "demand and supply are equal"? When people make this statement, what do they mean?

2. The demand and supply schedules for a certain grade of butter in the wholesale market of a certain city on a certain day are presumed to be as follows:

Prices	*Quantities that would be*	
Cents	*taken*	*offered*
10	14,000 pounds	6,500 pounds
11	12,000 pounds	7,000 pounds
12	10,500 pounds	7,500 pounds
13	9,000 pounds	8,000 pounds
14	8,000 pounds	8,500 pounds
15	7,000 pounds	9,000 pounds
16	6,000 pounds	9,500 pounds
17	5,500 pounds	10,000 pounds
18	5,000 pounds	10,500 pounds
19	4,500 pounds	11,000 pounds
20	4,000 pounds	11,500 pounds

(*a*) Draw the corresponding demand and supply curves.
(*b*) What is the price?
(*c*) In what respect does this problem differ from the preceding?

3. Referring to the demand and supply conditions as presumed in exercise 2:
(*a*) What would happen if a buyer should offer to take a large quantity of this grade of butter at a price of sixteen cents?
(*b*) What would happen if a seller should offer to sell at fifteen cents?
(*c*) What would happen if a seller should offer to sell at thirteen cents?
(*d*) What would happen if a buyer should offer to pay twelve cents?

4. Explain the difference between the two possible definitions of demand, illustrating by reference to the demand curve of exercise 1.

5. Explain the difference between the two possible definitions of supply, illustrating by reference to the supply schedule of exercise 2.

6. Using the definitions of demand and supply as adopted in this book, which of the following expressions is permissible and which is not permissible? Explain your answer in each case.
(*a*) The demand for automobiles exceeds the supply. no
(*b*) The demand for straw hats is less than usual this spring. O-K
(*c*) The greater the demand, the higher the price. yes
(*d*) Demand and supply are equal at the market price. [illegible]
(*e*) A drouth in the Middle West will decrease the supply of wheat. No
(*f*) A change in fashion may increase the demand for white shoes. yes
(*g*) The higher the price, the less the demand. No

7. Read the market reports in any newspaper and mark the words "demand" and "supply" wherever they occur. Try to frame a definition to fit the word as used in each case, noting especially whether the words are used according to the definitions adopted in this book.

CHAPTER XIV

THE CAUSES OF CHANGE IN COMPETITIVE PRICE

What causes price changes? Our excursion into the theory of price has thus far shown us the way in which competitive market prices are determined, as of a particular time, by demand and supply or by the decision of a sole buyer. Prices in the practical world of today are however seldom stationary, but are generally subject to change, more or less frequent. This well-known fact was specifically brought to our attention in Chapter XI. It is not enough therefore to know how a certain competitive price may be established in a static market. We must also understand the mechanism by which a price may rise and fall under the dynamic conditions of modern life.

A glance at the diagram in Figure 10 of the previous chapter will show that no other price than *pb* is possible with those particular curves of demand and supply, since *p* is the only point at which those curves can possibly intersect. A different point of intersection could be obtained only by a shift in one or both of the curves. In general terms a price which is determined by demand and supply can undergo change only as the result of a change in demand or a change in supply or simultaneous changes in both demand and supply; the same is true of the corresponding quantity exchanged. This indicates the path of our present inquiry.

Effect of an increase in demand. Let us continue to make use of our example of the egg market of A. The diagram in Figure 10 of Chapter XIII represents conditions as they are presumed to have been on October 15, 1924. A certain demand and a certain supply gave a price of sixty cents. Now let us suppose that there has been an increase in the demand for eggs and that on December 15 the demand is as shown in the following schedule:

DEMAND FOR EGGS. RETAIL MARKET OF A. DECEMBER 15, 1924

Prices	*Quantities that would be taken*
$1.00	275 dozen
.95	285 dozen
.90	300 dozen
.85	320 dozen
.80	340 dozen
.75	360 dozen
.70	380 dozen
.65	400 dozen
.60	420 dozen
.55	440 dozen
.50	470 dozen
.45	500 dozen
.40	540 dozen
.35	580 dozen
.30	630 dozen
.25	680 dozen

The reader is again reminded that the exact quantities in such a schedule are not significant but are used only for purposes of illustration. What we have is a new schedule in which, for each price, the quantity that buyers would choose to take is greater than before. This, it will be noted, agrees with the concept of demand

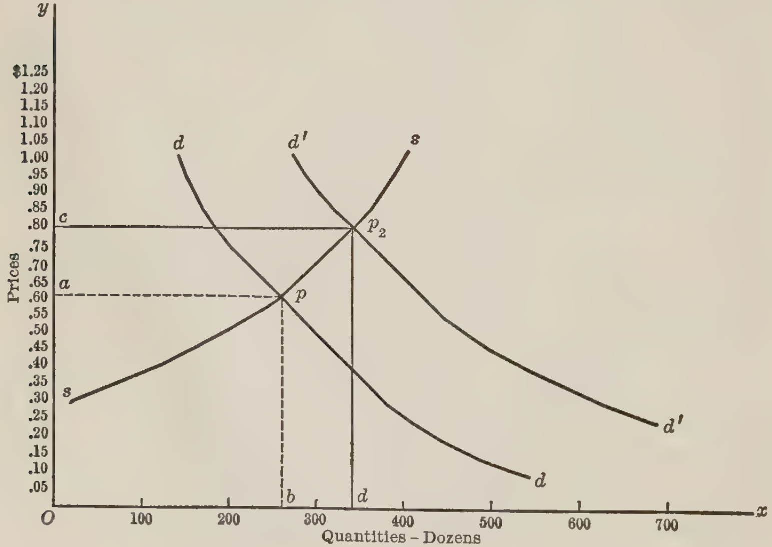


FIG. 16.—Retail Egg Market of A. December 15, 1924.

as we have defined the term. In common language also this is exactly what would be meant if for example people in A should say: "Cold weather is coming on and people are consuming more eggs; there is therefore an increased demand for eggs."

In Figure 16 on page 264 we have repeated the demand and supply curves as of October 15 (see Figure 10, Chapter XIII) and have added the curve representing the demand on December 15, as plotted from the schedule of the demand on that date. It will be evident that an increase in demand is represented graphically by a new curve, each of whose abscissas is longer than before; in other words, the demand curve has been shifted to the right. The increased abscissas indicate that now, at each price, the quantity that buyers would choose to take is greater than before. This again is exactly what is meant by an increase in demand. It is what people mean when they say: "The demand for eggs has increased since October"; "There is a strong demand for eggs today," and such expressions. An increase in demand then is always indicated graphically by a new demand curve further to the right.

What are the effects of the new demand? Let us assume for the sake of simplicity that there has been no change in the conditions of supply in the egg market of A. The curve ss (Figure 16), representing the supply on December 15, is then the same as for October 15. The curve $d'd'$ represents the demand on December 15. This new demand curve intersects the supply curve at the point p_2, further to the right and further up than the point p, which determined the price on the previous date. This means that there is a new price and a different quantity exchanged. The former price (October 15) was sixty cents, determined by the intersection of the demand and supply curves at p. Now (December 15) the demand and supply curves intersect at p_2, which means that at a price of eighty cents (measured by the line p_2d) the quantity taken and the quantity offered are the same; *i.e.*, 340 dozen (measured by the line p_2c). The price is therefore eighty cents, and the quantity of eggs exchanged is 340 dozen, an increase in both magnitudes since the earlier date. We conclude then that under conditions such as these an increase in demand tends to cause an increase in price and an increase in the quantity exchanged.

The reader will easily verify this conclusion by reference to his experience or his knowledge of ordinary price changes. For example he will perhaps recall the period of great business prosperity in 1920, when many people found themselves in possession of unusually high incomes. One of the first ideas commonly associated with an increased income is the purchase of an automobile. The demand for automobiles in 1920 was extraordinarily high, and the high prices of automobiles in that year as well as the extraordinarily large sales are well remembered facts. A moment's thought will suffice to bring to mind other examples.

Effect of a decrease in demand. The method just employed in the study of increasing demand will enable us to discover the effects of a decrease in demand. Let us suppose that, while conditions of supply remain unchanged, a slight decrease in the demand for eggs in A has occurred between December 15, 1924, and January 15, 1925. The new demand may be assumed to be as shown in the following schedule:

Demand for Eggs. Retail Market of A. January 15, 1925

Prices	*Quantities that would be taken*
$1.00	235 dozen
.95	245 dozen
.90	260 dozen
.85	280 dozen
.80	300 dozen
.75	320 dozen
.70	340 dozen
.65	360 dozen
.60	380 dozen
.55	400 dozen
.50	430 dozen
.45	460 dozen
.40	500 dozen
.35	540 dozen
.30	590 dozen
.25	640 dozen

In Figure 17 the corresponding curve, $d''d''$, is plotted. In this diagram we have also the supply curve, ss, which we assume has still not changed, and, for the sake of comparison, the former demand curve, $d'd'$, as it was on December 15. The new demand curve, $d''d''$, is evidently further to the left than the former demand curve. The abscissas of the new curve, which measure the quantities that would be taken at the various prices, are all shorter. This means that at any given price the quantity that buyers would take is less than before, which is exactly what is meant by a decrease in demand. Graphically a decrease in demand is represented by a new demand curve further to the left than the previous curve.

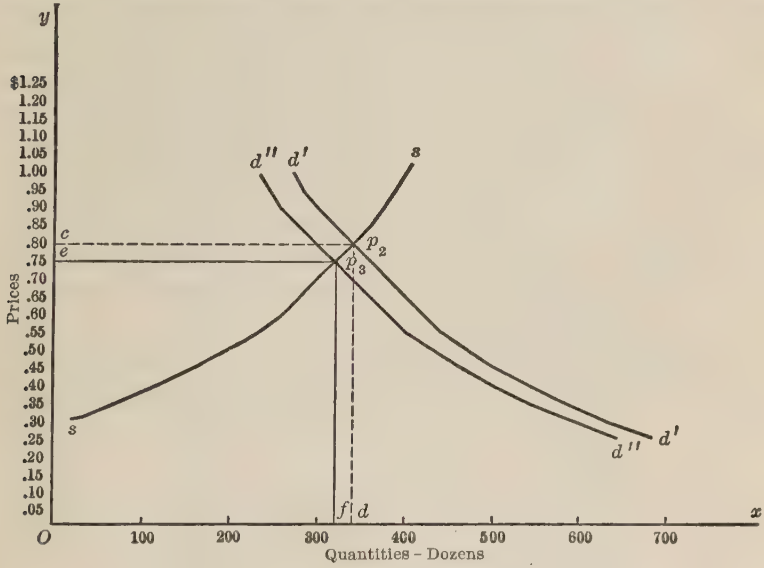


FIG. 17. — Retail Egg Market of A. January 15, 1925.

The new demand curve intersects the supply curve at a new point, p_3, further to the left and lower down than the former point of intersection, p_2. This indicates that there is a new price. The intersection of the curves at p_3 shows that at a price of seventy-five cents (measured by the line $p_3 f$) the quantity that buyers would take and the quantity that sellers would offer are each 320 dozen (measured by the line $p_3 e$). The price is therefore seventy-five cents and the quantity of eggs exchanged is 320 dozen. The generalization is that under these conditions a decrease in demand tends to cause a decrease in price and a decrease in the quantity exchanged.

Everyday illustrations will occur to the reader at once. Compare two years, in the first of which there was an early spring, while the second is marked by cold weather till late in the season. People in the second year will not want their usual quantity of spring suits, dresses, waists, and other seasonal goods. The demand will be small as compared with the year before and the result will be both a lower scale of prices and a smaller quantity of goods sold than in the normal season. An unusually warm winter will frequently cause a decline in the demand for furs, with the result of smaller sales and distinctly lower prices. Changes in fashion may suddenly check the demand for some article, following which there is sure to be a smaller quantity sold and a fall in price, if other things remain equal.

Effect of an increase in supply. Price and quantity exchanged are also affected by changes in supply. Our example of the egg market

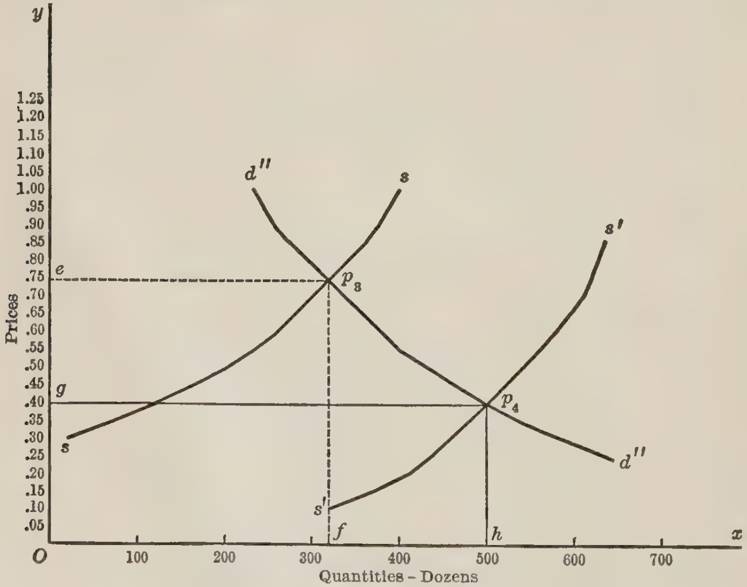


FIG. 18. — Retail Egg Market A. May 1, 1925.

of A will still serve us, by now illustrating the effects of increases and decreases of supply. The months of October to January are a period during which the average hen is at a low stage in her egg-laying activity. From March to June things are going much more briskly. An investigation carried out by the Connecticut Agricultural College showed that the average production per hen at this season was nearly three times what it was at the low stage in October to January. This increased production has an effect on the supply of eggs, which, as everyone knows, is normally much greater in the spring and summer than in the late fall and winter. Let us suppose the supply of eggs on May 1, 1925, in the market of A to be as shown in the following schedule:

Supply of Eggs. Retail Market of A. May 1, 1925

Prices	*Quantities that would be offered*
$0.85	635 dozen
.80	630 dozen
.75	620 dozen
.70	610 dozen
.65	595 dozen
.60	580 dozen
.55	560 dozen
.50	540 dozen
.45	520 dozen
.40	500 dozen
.35	480 dozen
.30	460 dozen
.25	440 dozen
.20	410 dozen
.15	370 dozen
.10	320 dozen

Let us assume that there has been no change in demand since the date of our last illustration, January 15. In Figure 18, we show the demand curve, *d''d''*, and the supply curve, *s's'*, as on May 1 and, for the sake of comparison, the supply curve, *ss*, as it was on the previous date, January 15.

There has been an increase in the supply of eggs. This means that, at any given price, the quantity that would be offered by sellers is greater than before. It appears from a comparison of the supply schedules of January 15 and May 1 and is represented graphically by a new supply curve further to the right than the previous one, with abscissas (measuring quantities offered) longer than those of the previous supply curve for each price. This agrees with our definition of supply. It agrees also with common usage in such statements as: "The hens are laying better and therefore there is an increased supply of eggs"; "The supply of eggs is very great just now," etc.

The new supply curve intersects the demand curve at a new point, p_4, indicating that at a price of forty cents (measured by the line p_4h) the number of eggs that would be offered is the same as the number that buyers would choose to take; *i.e.*, 500 dozen (measured by p_4g). The price is therefore forty cents, and the quantity exchanged is 500 dozen. The price is evidently lower and the quantity exchanged greater than before, and we have the general conclusion that, under such conditions as are here assumed, an increase in supply tends to cause a decrease in price and an increase in the quantity exchanged.

The harvests furnish a good illustration of this generalization. During the growing season the wheat market receives frequent reports as to the area planted to wheat, the progress of the crop, weather, rainfall, and other conditions affecting the future harvest. On the strength of these reports, sellers determine what quantities they can offer at various prices; that is, they determine the supply. A succession of favorable reports will thus increase the supply on the market, whereupon the price will fall and the amount sold will increase. Every observer of the market will recognize this relation between supply and price and quantity exchanged. Following the high prices of 1920, there came a distinct fall in automobile prices a few years later. There were other causes, but one of the most important was the increased supply offered by the manufacturers and dealers.

Everybody knows that prices in the city meat and fresh vegetable markets go tumbling down on Saturday afternoon and evening. Why is this? It is certainly not on account of a decrease in demand. Housewives are laying in their supplies for over Sunday, and the demand is, if anything, greater than on other days. The reason is that the dealers, knowing that much of their stocks will not keep till Monday, are offering to sell at greatly reduced prices. Stated in terms of supply, at any given price they are offering greater quantities than earlier in the week. This means that the supply has increased; hence the fall in price and the brisk selling. This example also brings out the difference between supply and stock on hand, to which attention has been called (see Chapter XIII). The dealers' stocks on Saturday evening are not larger than usual; indeed, as the evening wears on they become smaller. Yet supply increases and the price falls.

Effect of a decrease in supply. The effect of a decrease in supply is similarly illustrated. We may suppose that the enthusiasm of the hens has somewhat abated by the middle of July, whereas there has been no change in the demand for eggs in the market of A. The new supply will be slightly less than on May 1 and may be assumed to be as in the following schedule:

SUPPLY OF EGGS. RETAIL MARKET OF A. JULY 15, 1925

Prices	*Quantities that would be offered*
$0.90	595 dozen
.85	590 dozen
.80	580 dozen
.75	570 dozen
.70	555 dozen
.65	540 dozen
.60	520 dozen
.55	500 dozen
.50	480 dozen
.45	460 dozen
.40	440 dozen
.35	420 dozen
.30	400 dozen
.25	370 dozen
.20	330 dozen
.15	280 dozen

In Figure 19 the corresponding curve, $s''s''$, is plotted. We have here also the demand curve, $d''d''$, assumed to be the same as on May 1, and also, for the sake of comparison, the former supply curve, $s's'$. There has been a decrease in the supply of eggs. At any price that may be selected the quantity that would be offered is less than before. This appears from a comparison of the schedules and is indicated by the shorter abscissas (measuring quantities

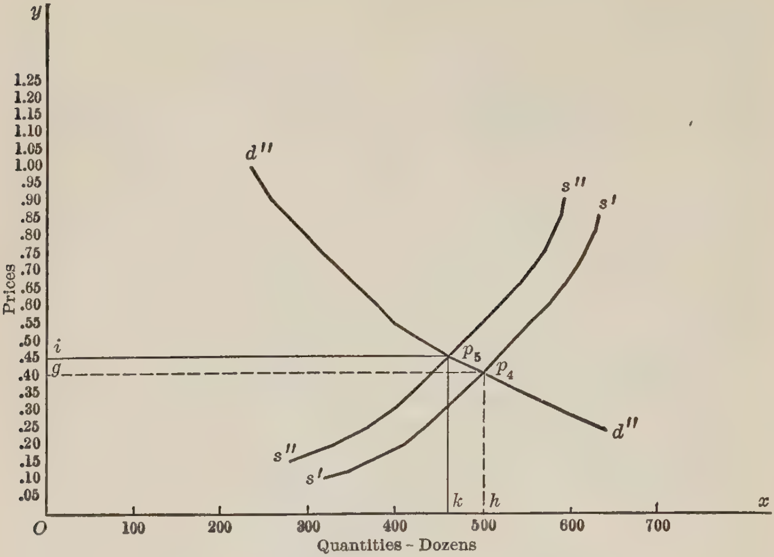


FIG. 19. — Retail Market of A. July 15, 1925.

offered) of the curve $s''s''$, as compared with the former curve $s's'$, owing to the position of the new curve further to the left.

The new supply curve intersects the demand curve at a new point, p_5, indicating that a price of forty-five cents (measured by p_5k) the quantity that would be offered, 460 dozen (measured by p_5i), is the same as the quantity that buyers would choose to take. The price therefore is forty-five cents and the quantity exchanged, 460 dozen. We have the generalization that, under such conditions, a decrease in supply tends to cause an increase in price and a decrease in quantity exchanged.

The harvests furnish practical illustrations of this principle, as they do of the effects of an increase in supply. Crop failures or adverse reports on crop conditions always reduce the supply, and, other things being equal, the result is a rise in price and a falling off in sales. The city of Santo Domingo depends for its stock of flour almost entirely upon imports. It happened one week during the World War that the arrivals of the regular ships were greatly delayed. The merchants' stocks were rapidly exhausted; the quantity obtainable from sellers at any given price rapidly declined. The supply decreased, the price of flour went soaring, and the quantity sold quickly declined. Any number of examples such as this will occur to the reader at once.

When the demand curve is vertical. The reader will not have failed to observe that the conclusions reached as to the effects of changes of demand and supply upon price and quantity exchanged were in direct consequence of the respective slopes of the demand and supply curves. Before drawing a general conclusion we must take account therefore of the peculiar condition of demand which gives rise to a vertical demand curve, as for example when a sole buyer calls upon competing sellers for bids to furnish a certain quantity of some article, as illustrated in Figure 20. Without following through the whole analysis in detail, it should be clear from a glance at the diagram that the vertical position of the demand curve makes no difference in the conclusions which we have previously reached, save for the one exception that changes in supply alone can obviously not cause any change in the quantity exchanged.

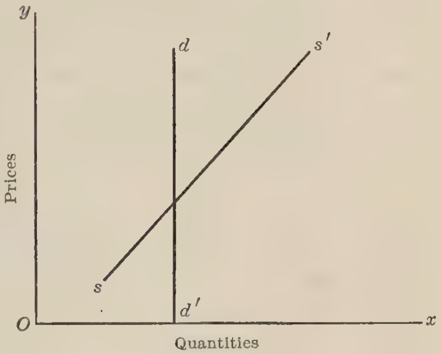


FIG. 20.

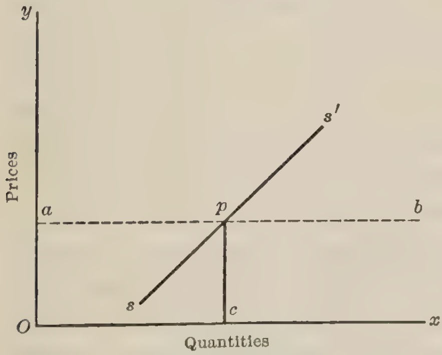


FIG. 21.

When price is fixed by a sole buyer. As regards the case of the sole buyer who fixes the price at which he will buy, there is no question of a change in price to be caused by changes in demand or supply, and there is no demand, in the economic sense, to undergo change. All that has to be considered is the effect of a change of supply upon the quantity exchanged, which will conform to the conclusion we have already reached, and the further observation that a change in the price offered by the buyer will cause a like change in the quantity sold, as illustrated by moving the line *ab* in Figure 21 up or down.

When cost of production fixes price. The situation here is much like that of the sole buyer discussed in the previous paragraph. There is no question of a change in price to be brought about by changes in demand or supply. There is no supply in the economic sense. All that we need note is that changes in demand will cause changes in quantity sold in harmony with our previous conclusion, and that if conditions of cost should cause a change in price (illustrated by a movement of the line *cc′* in Figure 22 up or down) an opposite change in the quantity exchanged would occur.

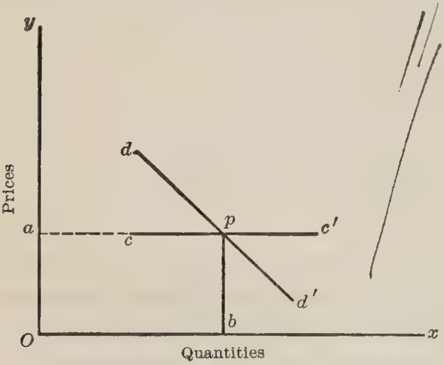


Fig. 22.

Conclusion. We have now investigated separately the effects upon price and upon the quantity exchanged of (1) an increase in demand, (2) a decrease in demand, (3) an increase in supply, and (4) a decrease in supply. The conclusions separately arrived at may be gathered together in a single generalization, as follows:

When there is competition of sellers, (1) a change in the magnitude of demand tends to cause a like change in price and a change in the magnitude of supply tends to cause an opposite change in price, except when price is fixed either by the decision of a sole buyer or by cost of production; (2) a change in the magnitude of either demand or supply tends to cause a like change in the quantity exchanged, except when the quantity is fixed by a sole buyer.

When demand and supply both change. This law covers all possible changes in either demand or supply. It is stated as a tendency, meaning that the results indicated will follow the respective causes unless some other cause interferes; that is, if other things are equal. In the illustrations thus far used we have assumed that other things were equal and have tested all possible changes of demand or of supply, the other factor remaining unchanged. But it is quite possible and normal that changes in both factors occur at the same time. The hens may lay more eggs, thus increasing the supply, at the same time that the approach of hot weather causes a decline in demand. Or both demand and supply may either increase or decrease simultaneously. In such cases the effect is the resultant of the two influences, as shown by the law of price and quantity exchanged. We must now test the law with respect to all possible combinations of changes in demand and supply.

Increased demand; decreased supply: effect on price. Suppose for example that, starting with our egg market of A on the date of our last illustration, July 15, the demand should increase while the supply decreased. The tendency of each change would be to increase price. Their combined effect would therefore be an increase in price greater than would have been caused by either change alone.

Let the reader refer to Figure 23. At the start the price was forty-five cents, determined by the intersection at *p* of the demand curve, *dd*, and the supply curve, *ss*. The new curves of demand and supply are $d'd'$ and $s's'$. If demand alone had changed, the intersection of the demand and supply curves would have been at *q*, determining a new price of fifty-five cents, as measured by *qi*. If supply alone had changed, the new intersection of the curves would have been at *r*, determining a new price of fifty-seven cents, as measured by *rf*. Both supply and demand curves having changed, their point of intersection is p'. This determines a new price, sixty-five cents, as measured by the line $p'h$. This price is obviously higher than the price before the changes in demand and supply, and also higher than the price that would have been determined by the change either in demand or in supply acting alone.

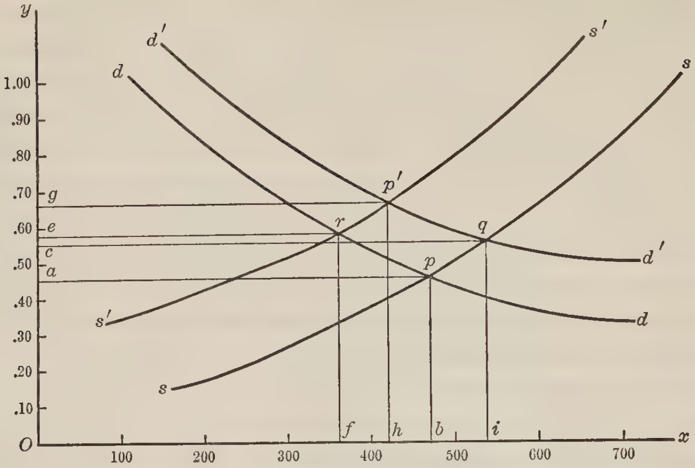


FIG. 23.

Effect on quantity exchanged. As regards the quantity exchanged the effects of these two changes, instead of supplementing, tend to neutralize each other. Referring again to Figure 23, we note that originally the quantity exchanged was *ob*, determined by the intersection of the demand curve, *dd*, and the supply curve, *ss*, at *p*. Now if demand alone had changed, the quantity would have been increased to *oi*; whereas if supply alone had changed, the quantity would have declined to *of*. The actual result, *oh*, is neither so great as *oi* nor so small as *of*, because the two effects have tended to offset each other. In this illustration the new quantity is less than the original, because the decrease in the supply was more important than the increase in demand. Another illustration might show the increase in demand overbalancing the decrease in supply and so causing a net increase in the quantity exchanged. Finally it is possible for the two changes exactly to neutralize each other and, while causing a rise in price, to leave the quantity exchanged as it was, as shown in Figure 24.

Decreased demand: increased supply. In like manner it may be shown that a decrease in demand and an increase in supply, tending each to lower the price, would in combination cause a decline in price greater than would have been accomplished by either

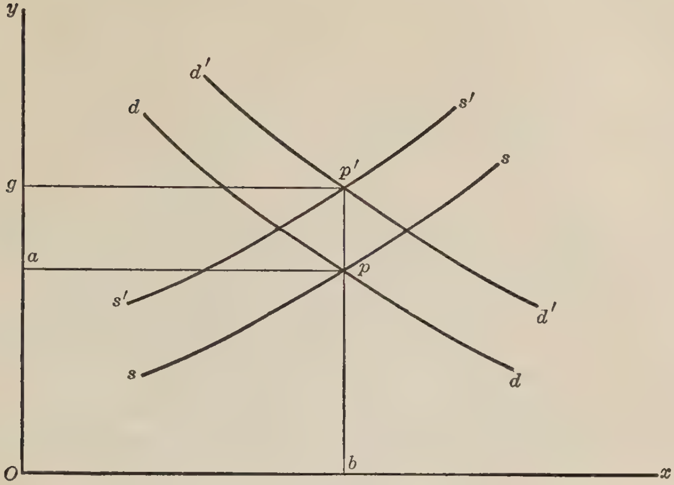


FIG. 24.

acting alone. And similarly it may be shown that a decrease in demand, tending alone to decrease the quantity exchanged, and an increase in supply, tending to an increased quantity, would in combination tend to neutralize each other and would cause either a decrease or an increase or no change at all in the quantity exchanged, depending on the relative influence of the two causes. The student should have no difficulty in demonstrating these propositions, following the method of the preceding paragraphs.

When demand and supply both increase: effect on price. As regards price, simultaneous increases in demand and supply tend to neutralize each other, since the tendency of the first is to raise price, while the second tends to lower price.

In Figure 25 we again take as our starting point the egg market of A on July 15, 1925. Now let us suppose the demand increases, giving a new demand curve, *d'd'*. The effect of this change in demand alone would be to raise the price to sixty cents, as measured by *qj*. But suppose that at the same time supply also increases. What will be the effect on price? Evidently the intersection of the new curves will have to be somewhere along the curve *d'd'* and to the right of *q*. The price therefore must be below sixty cents. In other words the increase in supply is tending to

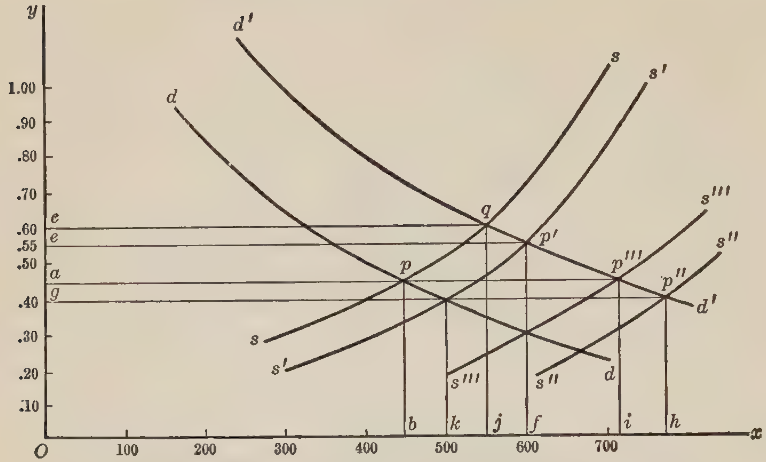


FIG. 25.

offset the increase in demand. Just where the price will be determined depends on the relative increases of supply and demand. If the increase in supply is relatively small, giving a new curve, $s's'$, the new price, as measured by $p'f$, will be greater than the original price (pb). If the increase in supply is relatively great, the new price may be lower than the original price. Thus a new supply curve, $s''s''$, gives a price, $p''h$, lower than pb. It is possible for increases in demand and supply exactly to neutralize each other and so cause no change in price, as would result from the new curves of demand and supply, $d'd'$ and $s'''s'''$, giving a price $p'''i$, equal to pb.

Simultaneous increases in demand and supply thus tend to neutralize each other, the resulting price being higher or lower than the original price, according to whether the predominant change is in demand or in supply, or being the same as the original price when the two changes exactly offset each other.

Effect on quantity exchanged. Since either of the causes now under consideration would, if acting alone, tend to an increase in the quantity exchanged, their combined effect must be an increase, and a greater increase than would have been caused by either alone. This is illustrated clearly enough in Figure 25. An increase in demand alone (from *dd* to *d'd'*) would have increased the quantity from *ob* to *oj*. An increase in supply alone (from *ss* to *s's'*) would have raised the quantity to *ok*. The two causes operating together make the quantity still greater, either *of*, *oi*, or *oh*, according to the assumption regarding the new supply curve.

When demand and supply both decrease. In like manner it may be shown that simultaneous decreases in demand and supply tend, as regards price, to neutralize each other, the resulting price being higher or lower than the original price, according to whether the predominant change is in supply or in demand, or being the same as the original price when the two changes exactly offset each other. On the other hand, either of these causes alone would tend to a fall in the quantity exchanged. The two together must therefore cause a fall in the quantity exchanged, and a greater fall than would have resulted from either change alone.

Other cases of competitive price. This analysis of the results of simultaneous changes in demand and supply has been confined to competitive prices determined by demand and supply of the ordinary kind. Corresponding conclusions would follow, were the analysis carried on to cover also the case of the vertical demand curve and those cases where price is determined either by the decision of a sole buyer or by cost of production. The reader should find no difficulty in making this further analysis for himself.

Practical illustrations.[1] The operation of these forces in various combinations may be observed at all times in the world of business and everyday life. For example consider the present prices of genuine antique oriental rugs in the United States as compared with the prices that prevailed in the late years of the nineteenth century. It is common observation that the prices have increased enormously. The explanation must be sought in changes in demand or supply or both. It is well known that the demand for oriental rugs has increased greatly. Never before was there such appreciation of the beauty and utility of these rugs. Never before were so many Americans willing to pay good prices for them. The quantity of any particular grade of rug that could be sold at any given price is greater than ever before; that is, the demand is greater. On the other hand the supply has declined. The World War was a disturbing factor, and even before that makers and traders had been diverted to the production of cheap modern imitations of the genuine antiques. It has become increasingly difficult to obtain the real thing, and the supply has declined. These two changes, an increased demand and a decreased supply, have acted together to cause the increase in prices.

---

[1] The reader has been reminded (see Chapter XI) that there has been a remarkable increase of prices in general since the latter years of the nineteenth century. This is to be regarded as due to simultaneous changes in the demand and supply of all goods consequent upon a depreciation in the value of the monetary unit. This is a topic whose full investigation must be postponed to a later chapter, following our study of money and banking. For the present however the reader must be reminded that, when speaking of increases or decreases of prices of particular goods in the following pages of this chapter, we refer to changes over and above the changes in the general price level affecting all goods alike; *i.e.*, we eliminate the changes in the general price level.

The retail market for fur garments in midsummer, as compared with the winter market, furnishes a good illustration of the opposite combination of changes in demand and supply. Of course people do not wear furs to any great extent in summer. Only those buyers will be in the market who are foresightedly providing for their coming winter needs. Both the number of the buyers and their eagerness to buy will be less than it was in the winter. In other words, the demand will be less. How is it with supply? The dealers are anxious to stimulate business in their dull season; they want to keep their store and office forces busy; they want to keep their capital turning over; they need money to meet current expenses. For such reasons they are willing to sell more cheaply than in winter. In other words, at any given price the quantity for sale will be greater; that is, there is an increased supply. The combination of decreased demand and increased supply is the sufficient explanation of the low prices which reward the thrifty summer buyer of next winter's furs.

About the time of the World War it rather suddenly became the fashion among the younger men of the United States to wear soft white shirts with collars attached. The demand for such shirts increased greatly. Other things being equal, the natural result would have been a great rise in prices. But other things did not remain equal. No sooner was the change in fashion evident than the manufacturers and dealers devoted themselves to the task of furnishing this type of shirt. Thus the supply increased, keeping pace with the increased demand. These two forces have thus counteracted each other, with the result that there has been no material change in the price of these shirts as compared with other styles.

Within a remarkably short lapse of time the automobile has virtually replaced the horse for pleasure driving and only slightly less completely for business purposes. The demand for harness for driving horses in the United States is only a fraction of what it was, say at the beginning of the present century. Has this decline in demand caused a fall in the prices of harness? Evidently not. A harness today costs no less than it did in 1900. This result, in the face of the virtual disappearance of demand, sometimes seems a strange thing. We have the answer in the law of price. A change in demand may be neutralized by a change in supply. With the advent of the automobile, the making of harness for driving horses declined. The supply declined and thus offset the decline in demand.

These examples will serve equally well to illustrate the effects of changes in demand and supply upon the quantity exchanged. Thus in the case of the white shirts with collars attached, the quantity exchanged has increased very greatly, which is the result to be expected from simultaneous increases in both demand and supply. The great reduction in the number of driving harnesses sold illustrates the effect of declining demand accompanying a decline in supply. The market for oriental rugs and the summer fur market are examples of changes in demand and supply whose effects upon the quantity sold tend to offset each other. Other examples will readily occur to the reader.

The most perfect examples of changes in demand and supply which neutralize each other arise in connection with variations in the value of money. This subject will be investigated in a later chapter. The reader is now merely reminded that he will later find, in the study of money and prices, additional material having an important bearing on price changes, acting always however through the media of demand and supply and cost of production.

EXERCISES

(Referring to pages 263 to 274 of the text.)

1. Refer to exercise 1 of the previous chapter (Chapter XIII). What circumstances might cause the demand for sugar to increase?
   (*a*) Draw up a new demand schedule indicating an increased demand.
   (*b*) Draw the demand curve corresponding to this schedule and, on the same diagram, the demand and supply curves of exercise 1 of the previous chapter (Chapter XIII).
   (*c*) What is now the price of sugar? How does it compare with the former price?
   (*d*) What quantity of sugar is now exchanged? How does it compare with the former quantity?

2. Refer again to exercise 1 of the previous chapter (Chapter XIII). What circumstances might cause the demand for sugar to decrease?
   (*a*) Draw up a new demand schedule, indicating a decreased demand.
   (*b*) Draw the demand curve corresponding to this schedule and, on the same diagram, the demand and supply curves of exercise 1 of the previous chapter (Chapter XIII).
   (*c*) What is now the price of sugar? How does it compare with the former price?
   (*d*) What is the quantity of sugar now exchanged? How does it compare with the former quantity?

3. State and solve a problem exactly like exercise 1 above, except that it deals with an increase in supply.

4. State and solve a problem exactly like exercise 2 above, except that it deals with a decrease in supply.

5. Note, from your own experience or knowledge, two or three illustrations of the effect upon price of each of the following causes: (*a*) an increase in demand, (*b*) a decrease in demand, (*c*) an increase in supply, (*d*) a decrease in supply.

(Referring to pages 275 to 282 of the text.)

6. Under what circumstances would a change in supply alone fail to make any change (*a*) in the price, (*b*) in the quantity exchanged? What would be affected by the change in supply in each of these cases? Draw diagrams to illustrate this problem.

7. Under what circumstances would a change in demand alone fail to make any change in the price? What would be affected by the change in demand in such case? Draw diagrams to illustrate this problem.

8. The demand and supply schedules for cotton in the cotton market of a certain city on a certain day are presumed to be as follows:

ordinates

Prices	*Quantities that would be*	
Cents	*taken*	*offered*
15	9,000,000 pounds	3,600,000 pounds
16	7,800,000 pounds	4,200,000 pounds
17	7,000,000 pounds	4,800,000 pounds
18	6,200,000 pounds	5,200,000 pounds
19	5,600,000 pounds	5,600,000 pounds
20	5,000,000 pounds	6,000,000 pounds
21	4,400,000 pounds	6,200,000 pounds
22	4,000,000 pounds	6,400,000 pounds
23	3,600,000 pounds	6,600,000 pounds
24	3,200,000 pounds	6,800,000 pounds
25	3,000,000 pounds	7,000,000 pounds

(*a*) Draw the corresponding demand and supply curves.
(*b*) Draw, on the same diagram, a new demand curve indicating an increased demand and a new supply curve indicating a decreased supply.
(*c*) What is the new price of cotton? Explain the change.
(*d*) What is now the quantity exchanged? Explain the change.

9. (*a*) Draw again the demand and supply curves corresponding to the schedules of exercise 8 above.
(*b*) Draw, on the same diagram, a new demand curve indicating a decreased demand and a new supply curve indicating an increased supply.
(*c*) What is the new price of cotton? Explain the change.
(*d*) What is now the quantity exchanged? Explain the change.

10. (*a*) Draw again the demand and supply curves corresponding to the schedules of exercise 8 above.
(*b*) Draw, on the same diagram, new curves indicating an increased demand and an increased supply.
(*c*) What is the new price of cotton? Explain the change.
(*d*) What is now the quantity exchanged? Explain the change.

11. (*a*) Draw again the demand and supply curve corresponding to the schedules in exercise 8 above.
(*b*) Draw, on the same diagram, new curves indicating a decreased demand and a decreased supply.
(*c*) What is the new price of cotton? Explain the change.
(*d*) What is now the quantity exchanged? Explain the change.

12. What changes in demand alone or in supply alone or in demand and supply simultaneously would cause:
(*a*) an increase in price?
(*b*) a decline in price?
(*c*) an increase in quantity exchanged?
(*d*) a decrease in quantity exchanged?

13. What simultaneous changes in demand and supply are possible without causing (*a*) any change in price? (*b*) any change in quantity exchanged?

14. What changes would you say had taken place in the last twenty years in the demand and supply of golf balls? How have these changes affected the prices of golf balls and the number of golf balls purchased?

CHAPTER XV

THE LAWS OF COMPETITIVE PRICE: SUMMARY, QUALIFICATIONS, AND CONCLUSIONS

We may now bring together, for convenient reference, the most important definitions and laws which have developed thus far from our study of demand and supply and price. This body of principles we may call, for short, the laws of competitive price.

The laws of competitive price. Summary: 1. *Definition of demand.* The demand for any good is a schedule of the respective quantities of that good which buyers would choose to take at all possible prices.

2. *Definition of supply.* The supply of any good is a schedule of the respective quantities of that good which sellers would offer at all possible prices.

3. *The law of demand.* In a given market at a given time, the quantity of any good which buyers would choose to take, when not a fixed quantity, varies inversely with the price.

4. *The law of supply.* In a given market at a given time, the quantity of any good which would be offered for sale, when not a fixed quantity, generally varies directly with the price.

5. *The law of competitive price.* (*a*) Competitive market price is determined (1) by demand and supply or (2) by the decision of a sole buyer or (3) by cost of production. In the first case the price is determined by demand and supply at that point at which the quantity which buyers would choose to take is equal to the quantity offered by sellers. The same forces which determine competitive price also determine the quantity exchanged.

(*b*) A change in the magnitude of demand tends to cause a like change in price and a change in the magnitude of supply tends to cause an opposite change in price, except when price is fixed either by the decision of a sole buyer or by cost of production.

(*c*) A change in the magnitude of either demand or supply tends to cause a like change in the quantity exchanged, except when the quantity is fixed by a sole buyer.

" Theory and practice." A scientific law is a statement of the necessary relation between things, or a statement of the effect which always tends to follow a certain cause. Every scientific law is thus a statement of a tendency; a statement that a certain result would always follow if certain conditions were fulfilled. If the conditions are not precisely fulfilled, the result does not necessarily follow. In the demonstration and statement of a scientific law the conditions may be assumed with a precision and minuteness which can never be obtained in actual practice. There is thus always a certain discrepancy between the result as stated in the law and the way things work in practice. This is true of all sciences which deal with practical affairs. The science of physics contains a body of laws whose precise statements never agree exactly with the way things work for the civil engineer who has a railroad bridge to build. The practical chemist must always expect a certain difference between the actual results he gets and the precise statements of his formulae. This does not mean that the scientific laws are false or inapplicable. Such discrepancies are due to the fact that it is never possible in the statement of a scientific law to take account of all the forces that will be found acting in the practical case. Some of the more exact sciences, such as physics, chemistry, and astronomy, are able to go far toward this ideal. Yet there is always something in the practical result that could not be exactly foreseen and provided for in the formula. A second cause of discrepancy is the fact that scientific laws can deal with the smallest conceivable magnitudes, infinitesimals as they are called, while in the practical problem the quantities handled, however minute, must have actual size.

The science of economics is especially subject to these discrepancies between principles and practical results, and we cannot safely neglect the task of examining our laws of competitive price from this point of view.

" Market price " and the " perfect market." First of all it is evident that the laws of price which we have developed do not explain every individual price in a competitive market. Attention has already been called to the fact that irregular scattering prices are likely to occur before the market settles down to equilibrium and that isolated irregular prices may occur at any time. A "perfect market" would be one in which all persons interested in either buying or selling the good were in uninterrupted communication with each other and were possessed of infinite knowledge of all the factors affecting demand and supply. In that sort of market, the true market price would be instantly established and never departed from, and our laws of price would admit no margin of error. But of course no such market exists, although the closeness of the approach in some of the organized markets for dealing in corporation securities and certain staple commodities would surprise the uninitiated. Our laws relate to the equilibrium, or prevailing, or "market" price; they do not deny that goods are exchanged from time to time at other prices, for whose variations there may be insufficient explanation.

The student must also remember that for any good there will generally be, not one market and one market price, but many markets and many market prices, each one the result of forces of the same general character but of different magnitudes in the several markets. The prices of a given good in different markets will seldom be wholly unrelated to each other, as has been already suggested, but nevertheless they are distinct prices.

The "smooth" curve. The thing that is most likely to mislead the student is not so much the statement of the laws of price as the process by which these laws have been developed and illustrated. Let us take for example the law of demand. We have stated it thus: in a given market at a given time the quantity of any good which buyers would choose to take varies inversely with the price. Now this principle is so constantly seen working in practical affairs, it is so much a matter of common knowledge, that its demonstration may have seemed hardly necessary. Anybody will admit its truth at once. But the curves by means of which we have illustrated it may have given to this law an appearance of

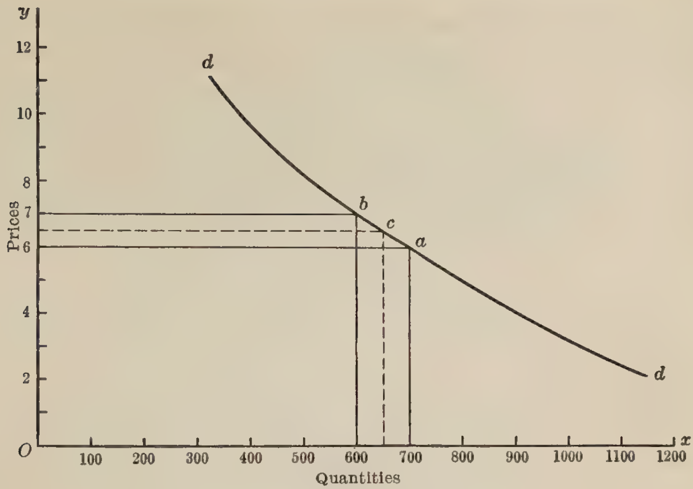


FIG. 26.

minute precision which it does not have in actual practice. The demand curve in Figure 26 would seem to indicate that the most minute change in price calls forth a correspondingly minute change in the quantity that buyers would take. If at six cents 700 units would be taken, as indicated by the point *a*, the next point on the curve above *a* (call it *a′*), separated from *a* by a distance infinitely small, should indicate that at a price only minutely above six cents a quantity minutely less than 700 would be taken. Mathematically there is an infinite number of points on the curve between *a* and *b*, whereas *b* may be the closest point above *a* which has any practical significance, simply because prices go up by one cent increments and there is no price between six cents and seven cents. If prices could vary by infinitely small amounts and quantities could be split up into infinitely small particles, a point *c* on the curve might show that at a price of 6.53217 cents the quantity (of sugar, for example) that would be taken would be 643 pounds and $749\frac{1}{5}$ grains. Of course this is of no practical significance. In making practical application, we must consider the curve not as a continuous curve, but as a series of separate points, *a*, *b*, etc., the parts of the curve between such significant points having no practical meaning.

Qualifications to the laws of price. Competitive price and the quantity exchanged are under certain circumstances the result of the joint action of demand and supply. The truth of this law cannot be questioned. But we must not be misled by the appearance of minute precision given by the curves which have been used to demonstrate it. Let Figure 27 relate to the retail sugar market of a certain place on a certain day, and suppose the intersection of the curves at p to indicate a price of 7.1 cents and a quantity exchanged of 400 pounds. Demand and supply curves, being independent of each other, may of course intersect at points giving fractional prices and fractional quantities. But actually there would not be a price of 7.1 cents. The price would be 7 cents. At that price the quantity offered, *ab*, is slightly less than the quantity that would be taken, *ac*. This discrepancy cannot be corrected by a rise in price, since retail prices of sugar are not quoted in small fractions of a cent. There will actually be then some buyers who will not get all that they want at seven cents, since the quantity exchanged cannot be greater than *ab*, the amount which sellers

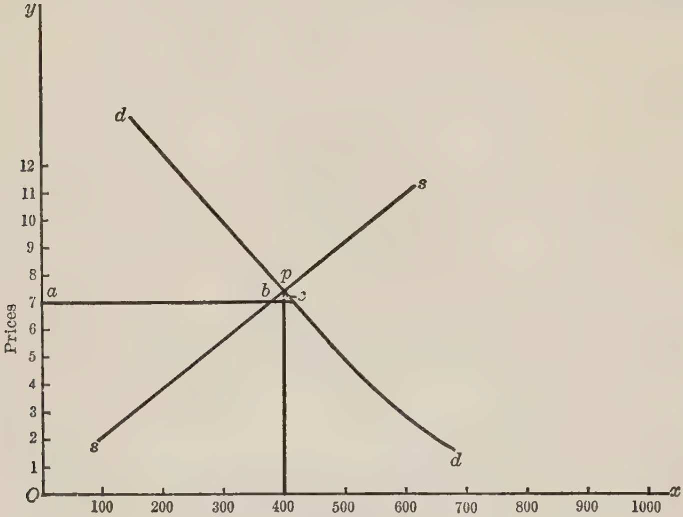


FIG. 27.

are willing to dispose of at seven cents. Similarly a different pair of curves might have led to a slight discrepancy the other way. At the actual price sellers might be glad to part with a larger quantity than buyers will take, but they cannot make the normal correction by lowering the price a full cent or half cent, since that would throw the market into still greater lack of equilibrium. We are thus compelled to qualify to this extent our previous general statements that, at the price as determined by demand and supply, (1) the quantities offered and that would be taken are equal, (2) these quantities could not be equal at any other price, and (3) this price clears the market.

We must make similar qualification to our conclusions as to the effects of changes in demand and supply. The laws as stated are true. But it is of course possible that a minute change in demand or in supply may not affect the price, simply because it is not enough to raise or lower the price by a whole cent or whatever is the smallest amount by which prices are actually changed. The price of milk (see Figure 28) may be ten cents and the quantity sold 600 quarts. In spite of the appearance of the curves, a minute increase in demand may not raise the price, simply because it is not enough to bring the price up to eleven cents or at least to ten

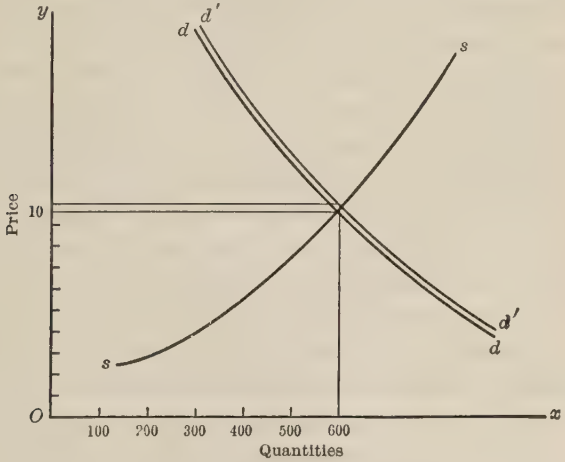


FIG. 28.

and one half cents. Without raising the price, an increase in demand will often increase the quantity sold, since sellers were previously willing to sell more than buyers would take at the price. A slight decrease in demand will sometimes decrease the sales, without changing the price, sellers having not yet reached the point where they are willing to make a material reduction in price, preferring rather to have some goods left on their hands. Similarly there may be slight changes in supply without causing changes in price, though they may increase or decrease slightly the quantity sold.

Cause and effect in practice. The action of the laws of price then is not smooth and instantaneous. Prices do not constantly oscillate back and forth like the sensitive magnetic needle on a storm-tossed ship. A considerable time may elapse before a change in demand or in supply works out its effect in price and quantity sold. Small changes may have no effect. Since conditions are constantly changing, perfect equilibrium is seldom reached and never long continued. Even the tendency toward equilibrium is not smooth, regular, and constant, like the swinging of the pendulum back and forth across the vertical position or the oscillation of the magnetic needle. The economic adjustment of price and quantity exchanged is more or less irregular; it moves jerkily by fits and starts. An analogy may be found in the operation of any machine. Everyone knows that the speed of an automobile, other things being equal, varies directly with the quantity of gas delivered to the motor. A smooth curve might be drawn showing this relation for any particular car. Yet in actual operation the correspondence between gas and speed will depart considerably from the precise continuous relation indicated by the curve. It will be impossible to detect changes in speed resulting from addition or subtraction of a small fraction of a drop of gasoline. An appreciable amount must be added or subtracted to get an appreciable effect. And even then the effect will not correspond exactly to the quantity of gas. There is play in the motor bearings and in the transmission gears. The traction of the tires on the road is not perfect and constant. The response of the car to the throttle will not be absolutely smooth and constant; it will be slightly jerky and irregular.

The laws of price on the stock exchange. Possibly the closest approach to a perfect market in the world is the New York Stock Exchange. The buying and selling of securities in this market furnishes excellent illustration of all the laws of price. It also illustrates the lack of perfect adjustment of cause and effect, even in the most nearly perfect market. Soon after the opening of the New York Stock Exchange at ten o'clock on the morning of July 28, 1924, 400 shares of the common stock of the United States Steel Corporation were sold at a price of $101\frac{7}{8}$. Of course the number of shares actually bought was equal to the number sold. But we cannot be sure that there were not some sellers willing to sell at this price but unable to find buyers. There may have been just one buyer calling for 400 shares and offering $101\frac{7}{8}$. If two or three brokers offered to sell him the 400 shares, he would pick one, presumably the one who called out his offer first. But wouldn't the others at once offer to sell at a lower price? Not necessarily. They may not be willing to take a lower price, preferring to wait in anticipation of another buyer or not to sell at all. Conversely one seller may have offered the 400 shares at $101\frac{7}{8}$ and found several buyers willing to take them at that price but unwilling to offer more. Furthermore dealing on the New York Stock Exchange is normally in 100 share lots, and this obviously prevents a smooth and continuous adjustment of quantity exchanged. So we have the price determined, as well as the quantity exchanged, but there is not the perfect "clearing of the market" which our smooth curves would seem to imply.

A little later on this morning the price of United States Steel stock dropped to $101\frac{3}{4}$, with a sale of 100 shares. Then it went back to $101\frac{7}{8}$, and 900 shares were sold at that price. So during the day the price oscillated back and forth between a high point of 104 and a low point of $101\frac{1}{4}$; the last sale of the day was at the highest price, 104. During the day 46,200 shares were sold (the reports take account only of lots of 100 shares or more). The price changed sixty-four times during the day, thirty-nine times in the last hour's business, between 2 and 3 p.m., when trading was very active. These are all facts, as reported on the stock ticker, in the commercial market reports, and in the New York and other newspapers of July 29, 1924. They show a high degree of adjustment between demand and supply and price and quantity exchanged, yet not necessarily a perfect adjustment.

Shall we discard our theory? There are those who lose all faith in the general principles or laws of economics when confronted with facts such as these. We hear often enough: "What is the good of all your theories if they don't agree with the facts?" or "That is all right in theory, but it won't work in practice." The careful student of any science will not fall into such hasty conclusions. He will realize the true relation between a scientific generalization and the observed facts. He will see that there is always a lack of perfect correspondence, a certain "margin of error." But he will not therefore throw his general principles overboard. He will simply seek to determine as closely as possible the margin of error and to make allowance for it. Our smooth curves are a very useful device for the investigation of the laws of price, even though they represent a smoothness, a constancy, and a precision which are not actually found in the market. Having learned how price would be determined under the ideal conditions implied in the curves, we are then ready to discover the discrepancies in practice, to explain them, and perhaps even to measure and allow for them. To throw overboard our economic generalizations and their graphical illustration, because they do not fit perfectly the actual facts, would be as foolish as for the automobile engineer to discard for the same reason his curves of the relation between quantity of gas and speed.

Above all the careful student will not make the foolish statement that some principle "is all right in theory but won't work in practice." Any generalization which "won't work in practice" is not sound theory. But lack of perfect correspondence between theory and practical facts does not mean that the theory will not work. In spite of all allowance for discrepancies, it cannot be doubted for a moment that demand and supply did determine the prices at which United States Steel stock was sold on the New York Stock Exchange on July 28, 1924. There can be no doubt that, in spite of lack of perfect adjustment, the quantity of a good offered for sale will generally vary directly with the price, and that the quantity which buyers will choose to take does, by and large, vary inversely with the price. Changes in demand and supply certainly do have the effects on price which are laid down in our laws. The laws are true, making due allowance for the margin of error. The laws are also practically useful. Hardheaded business men, the first to recognize their lack of perfect application, have always relied, consciously or unconsciously, on the economic laws of price, and the up-to-date business man is going to greater lengths than ever before in the application of economic theories to the practical conduct of his business.

The students of the physical sciences and the practical men who work in those fields have generally outgrown the crude notions about " theory and practice " which still make trouble in economics and business. One of the most important tasks before the student of economics is to learn the real meaning of economic laws, their proper limitations, and the way to apply them to practical problems.

Purpose of this study. We have now studied, in considerable detail, various aspects of demand and supply in their relation to the prices of goods and to the quantities of goods exchanged on the market. The reader should perhaps be warned that our purpose in all this was not to arrive at a set of rules to be memorized. The student who undertakes merely to memorize all of these conclusions from the study of demand and supply and price will have set himself a laborious and barren task. On the other hand the student should by now be able to derive for himself any of these conclusions whenever he needs it. The purpose of this study was not to develop certain formal rules for their own sake. It was rather to furnish the student with the mental equipment needed for the mastery of the economic problems of life. These problems in the modern world are many and complicated and difficult. The best equipment for meeting them is not a mind stored with formal rules mechanically memorized but a mind trained to reason accurately from given premises to correct conclusions. The student should test the success of his study of the foregoing chapters, not by seeing whether he can repeat from memory all the conclusions as to the operation of demand and supply, but by applying his knowledge of the nature of demand and supply to the exercises here presented and to other similar exercises which he can set for himself.

How to study problems of price. In particular the student should have a clear understanding of just what demand and supply are and a feeling of confidence in his knowledge. He should have mastered such use as has been made of schedules and curves, not at all for their own sake, but because they are the most convenient tools for the study of economic problems, bringing simplicity and clearness out of situations too complex for the mind readily to grasp without their aid. The student should acquire the habit, whenever he has to think of demand, supply, price, and problems involving them, of drawing on a piece of paper two lines meeting to form a right angle and, within the angle, two intersecting curves or lines, as in Figure 29. This will clarify any problem and start the investigator on the road to its solution. Finally the student should have been fully warned of the dangers of loose methods of thought and careless conclusions. He should realize that only by precision in his concepts and rigid accuracy in his reasoning can he achieve success as a student of economics or gain the mastery of the economic forces which so largely dominate the world in which we live.

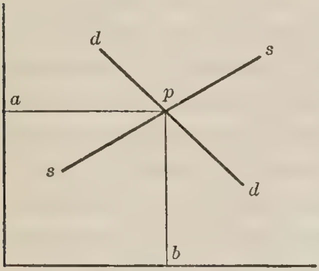


FIG. 29.

EXERCISES

1. Conditions of production determine the cost of a certain article at $5.00, at which price any quantity up to 100,000 will be offered by sellers. Draw a diagram illustrating the determination of the quantity exchanged under these conditions.

2. Changed conditions make it possible to produce the article considered in exercise 1 above at a cost of $4.00. Other things being equal, what will

be the effect of this change on the quantity exchanged? Illustrate by a diagram.

3. Explain, to the best of your knowledge or supposition, with illustrative diagrams, the following price comparisons, assuming also such changes in the quantities exchanged as would appear reasonable to you (disregarding any change in the general price level):

  (*a*) Pine boards cost twice as much in 1891 as in 1860.
  (*b*) The price of sugar in 1891 was about half its price in 1860.
  (*c*) The price of milk was the same in 1860 and 1891.

4. A manufacturer asks for bids to furnish 1,000 machines of a special kind for which there is no sale elsewhere. He is prepared to pay any price up to $150 apiece.

  (*a*) Draw supply and demand curves illustrating this situation. What price does he have to pay?
  (*b*) Can you draw supply and demand curves indicating a condition under which there could be no sale?

5. A fur dealer offers to pay $5.00 apiece for all skins of a certain kind brought to him.

  (*a*) Draw a diagram illustrating this situation.
  (*b*) According to your diagram, how many skins does he purchase? How many would he have purchased if he had set his price at $7.00? at $4.00?
  (*c*) Can you draw a diagram indicating a condition under which there could be no sale of skins to this dealer?

6. Look up the market reports in a recent newspaper and note the prices of some stocks or bonds or commodities about which you may have some special knowledge. Now look up the prices of the same securities or commodities in a newspaper some months or years earlier. How have the prices changed? Can you suggest the probable causes of these changes?

Hand these in -

CHAPTER XVI

WHAT IS BACK OF DEMAND

Direct and indirect determinants of price. In the foregoing chapters of Part II the reader has acquired some knowledge of the economic laws relating to competitive price. He has learned that competitive price is determined usually by demand and supply, though sometimes by the decision of a sole buyer or by cost of production, and by these causes alone. This may at first seem a hard saying. Is not the price of eggs affected by the activity of the hens in laying? Does not the rainfall affect the price of wheat? Does not a late, cold spring reduce the prices of straw hats, and an early, cold winter enhance the prices of furs? Are not the prices of many articles affected by changes in fashion? And are we not always being told that the "high cost of living" is due to "inflation of the currency"? Such pertinent questions might be multiplied indefinitely, but their obvious affirmative answers do not imply that here is a host of additional causes affecting competitive price. All these circumstances and hundreds of others do influence prices, but they do so only indirectly. They act only by first affecting either demand or supply, cost of production, or the decision of the occasional sole buyer. Our previous study has at least indicated this. A late, cold spring causes people to postpone or give up the purchase of the new straw hat; thus the demand is small and the price falls. A drought reduces the wheat harvest; the supply is diminished, and the price of wheat goes up. There is no side door through which any influence can get at price. Price may be approached only through one of the specified main entrances.

It follows however that in developing the so-called laws of price we have by no means solved the whole problem of price. There remains the important task of discovering how demand and supply and cost of production are themselves determined. Our pursuit of this investigation will introduce us to some of the most fundamental problems of economics. Turning first to the buyer's side of the market, let us undertake to discover what is back of demand.

Total demand and individual demand. The demand for any good in any market is the result of the attitude of all the prospective buyers (that is, all those actively interested in buying) in that market. There may be any number of these buyers, from one to hundreds or thousands. We have assumed certain schedules and curves representing the respective quantities which buyers would presumably choose to take at each possible price. Each of these quantities must be the total of the several amounts that the various individual buyers would choose to take at the corresponding price. Evidently the first step in analyzing demand is to break up the total demand (which is what has been considered thus far) into its component parts; *i.e.*, the individual demands of the several buyers. We must inquire why anyone wants to buy wealth and services, what determines the prices he is willing to pay, what decides the quantity of each different thing which he would choose to buy.

Of course a person buys anything because he desires it to satisfy a want. But whenever anyone buys something, he has to give money in exchange for it. He must therefore want the thing bought more than he wants the money[1] he has to give for it. But why does he stop buying? How does he decide how much to buy of any particular kind of wealth or service? To answer these questions we need to make further investigation into the subject of human wants. We have already learned something of the nature of man and of the environment in which he lives. We have learned that man has an infinite number and variety of wants. We must now give attention to the way in which any particular want is satisfied.

How utility varies. Suppose you are playing golf on a day when the clubhouse employees are off on a vacation and it is impossible to buy golf balls at the club. Your drive off the first tee goes into the rough, and the ball is lost. If that happens to be the only golf ball in your possession, your chagrin at its loss will be extreme, far greater than is measured by the price paid for that particular ball. Seeing your whole day's sport gone, you say, "I would give ten dollars for any old golf ball right now." Ten dollars then may be taken as the measure of your loss. Now let us change the supposition by assuming that your bag is packed with a dozen new balls. What does the loss of the first drive mean to you now? Probably nothing more than the money value of a ball. If, instead of a dozen, you had had half a dozen golf balls on which to rely for the day, the loss of one would have been a more serious matter. If you had started with only four, or three, or two, the loss would have been more serious still. Changing the illustration again, let us suppose that as you play through the course you meet a thrifty caddy who offers to sell a golf ball which he has found, but only at the highest possible price for which he can hold you up. Ignoring any restraining club rule, how do you decide what you will pay? The first thing is certainly to take count of your stock. If you happen to have just driven your last one into the woods, you will if necessary pay ten dollars or whatever sum measures the importance to you of the chance to continue the game. If you have just one golf ball left, your offer will not be so high, but it will still be high. If you have two, or three, or four, your offer will be correspondingly lower, till finally you may be so well supplied that you will decline to pay more than the regular price or even refuse to buy at all, virtuously deciding not to countenance the boy in his nefarious practice. We can readily draw the general conclusion that the utility of an additional golf ball varies greatly, depending at any given time upon the number of golf balls possessed, and varying inversely with the number possessed. The same rule holds of the utility of a golf ball subtracted from the stock on hand.

---

[1] Money of course belongs to the class of intermediate goods and is wanted not for its own sake but because of its power to give command over other goods, a point to which we shall return later in this chapter.

The subjective side of utility. This example will serve to introduce a human characteristic which relates itself to all goods (wealth and services). Heretofore we have studied utility objectively, as an attribute of wealth, that quality of wealth and of free persons which enables them to satisfy human wants. We must now consider utility subjectively, from the viewpoint of the person whose want may be satisfied. Subjectively utility expresses the estimation or valuation placed upon a service or upon the wealth rendering the service, by a person who desires that service. When a person possesses more than one unit of any good, the wants which are destined to be satisfied by the several units are not of equal importance. Suppose there is a water famine and a family finds itself in possession of ten buckets of water for the day's use. One bucket will be set aside for drinking, one for cooking, one or two for washing the dishes, two or three for bathing, three or four for washing clothes, and so on down the list of uses of water in a descending scale of importance. If the number of buckets on hand were less, some one or more of the less important uses would have to be given up; perhaps there could be no washing of clothes that day. If the quantity were greater, other uses still less important would appear — it might even be possible to water the lawn. In this example the wants to be satisfied by the several units are different in kind as well as in importance, but the principle holds equally when the wants are all of the same kind. The first piece of candy given to a child means a good deal to him. Give him another, and his appreciation is somewhat less. The second, third, and fourth mean successively less to him, and the point may finally be reached where he has no desire for another piece. This is the point of satiety. Even though the uses of the several pieces of candy are identical, the wants satisfied by the several pieces are of progressively diminishing importance.

Utility expressed quantitatively. In order to give a quantitative form to its investigation, we must have some way of measuring utility. The institutions of exchange and money are so completely established in the modern world that we are accustomed to measure our desires and our estimates of things in terms of money, even when there may be no question of buying and selling. It is convenient and usual therefore to measure utility in terms of money, and we may take as the unit for measuring utility the monetary unit.

With this yardstick we may construct a quantitative picture of marginal utility, by measuring the utility of one golf ball to our imaginary player under various assumptions as to the number he has on hand, continuing the assumption that he has no means of obtaining any more on this day. The result will be a schedule such as the following:

Quantity possessed	*Utility of one unit*[1]
1	$10.00
2	8.00
3	6.00
4	4.00
5	2.50
6	1.50
7	1.00
8	.75
9	.50
10	.25
11	.10
12	.05
13	.00

Marginal utility. We call this the schedule of *marginal utility*, and we define marginal utility as follows: *The marginal utility of any good to any person at any particular time is the utility to him of one unit of it.*

There

The curve of marginal utility. We may represent our golf player's mental attitude graphically as in Figure 30. The number of golf balls on hand is measured along the x axis. Each vertical ordinate (*na*, *pb*, etc.) measures the utility of one golf ball when the number possessed is as indicated on the x axis. Thus if he were down to his last golf ball, its utility would be ten dollars, as measured by *na*. If he had five, the utility of one would be two dollars and a half, as measured by *se*, and so on.

The reader needs to be warned against a possible error here.

---

[1] Although a golf player would not ordinarily have so many golf balls that one would be worth less to him than the regular price, we may nevertheless, for the sake of theoretical completeness, continue the schedule till the utility becomes zero, since it is conceivable that a player might be burdened with such a quantity of golf balls (with no other means of disposal) that he would give up one or more for little or nothing simply to be relieved of the necessity of carrying them.

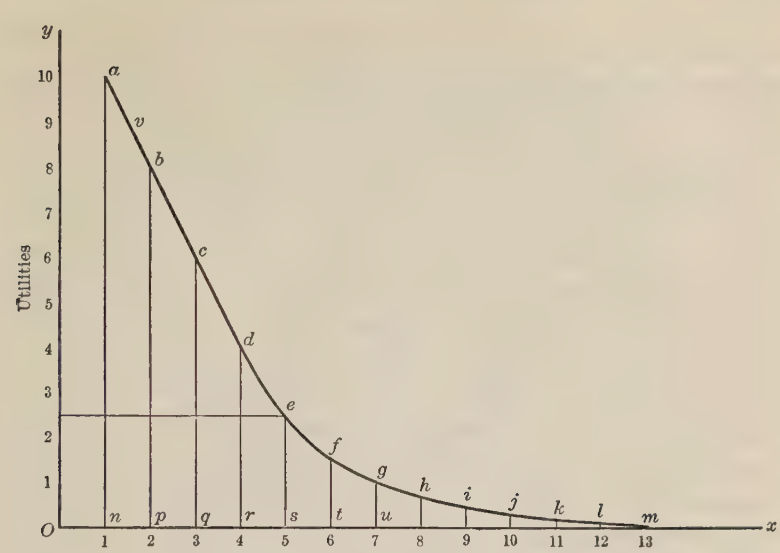


FIG. 30.

The points on the curve *am* (Figure 30) other than those marked by the letters *a*, *b*, *c* . . . *m*, have no significance. We can say that the marginal utility is ten when the number possessed is one, as indicated by the point *a*, or that it is eight when the number possessed is two, as indicated by the point *b*. But, since we cannot deal with fractions of a golf ball, an intermediate point, as *v*, means nothing. Strictly the diagram should not be a curve at all, but merely a series of ordinates, *na*, *pb*, *qc*, etc. We join the ends of the ordinates with a curve simply in order to bring out more clearly the relation between the marginal utility and the quantity of the good possessed. When it is possible to use very small units, the points may become so numerous as to give us practically a smooth curve. Thus if we represented the marginal utility of sugar, measured in pounds, the diagram would be somewhat as in Figure 30. Measured in quarter-pounds, we should have four times as many significant points. Measured in grains, the points would be so close to each other as to form practically a continuous curve. Actually we seldom are able to use units small enough for such a result.

Meaning of marginal utility. We must be clear as to exactly what marginal utility means. The curve does not show of course what was actually paid for the successive units or for any particular unit. It does not show a historical sequence of uses or utilities; it relates to one particular time and shows what, at that time, would be the utility of one unit if the possessor had any given number. It is somewhat misleading to say, as is sometimes done, that marginal utility is the utility of the last unit. The units are all alike. Having any given number, one will not be preferred over another; their utilities will all be the same. It is the uses to which they may be put that are of varying importance. The marginal utility is thus measured by the least important of the uses that would be resorted to when that particular number of units is available. Only in this sense can we speak of the "last" unit. The curve of marginal utility simply shows what, at any given time, would be the estimate of the least important use to which a unit would be put, assuming any given number of units to be possessed. Hence our notion of marginal utility as the utility of one unit of the stock on hand.

The law of marginal utility. Once we have a clear idea of the meaning of marginal utility, the economic law of the relation between quantity possessed and marginal utility, or *the law of marginal utility,* follows almost as a matter of course. It may be stated thus: *The marginal utility of anything to any person varies inversely with the quantity which he possesses.* A glance at Figure 30 will make the meaning of this perfectly clear. This law holds, as we have seen, equally whether the successive uses to which a good is put are apparently identical, like the bites of an apple, or are different, as the uses to which different buckets of water may be put.

The psychological foundation. This law is a result of the well-known psychological law that when an identical stimulus is repeated the response gradually becomes weaker. The appetite for food is the simplest illustration. Having eaten one apple, the second gives us less satisfaction, the third still less, and very soon we reach the point of satiety, when no further satisfaction, perhaps even repugnance, results. The same loss of appetite would have appeared more gradually if we had considered the separate bites. This principle will be recognized at once as holding good of the satisfaction of practically all wants, though there may appear to be exceptions when for a time the satisfaction seems to increase, as for example when one listens to music and, gradually getting into the spirit of it, gets more pleasure from the second piece than from the first. Whether such cases are true exceptions or not, the point of diminishing utility is sure to be reached sooner or later, which is the essential point for our purpose.

Regard for future wants. This psychological law is not alone sufficient to explain the law of marginal utility. When a man buys bread he does not determine its marginal utility simply on the basis of the satisfaction he could get from eating bread then and there. If so, he would place a high value on a single loaf; a second might have a slight utility if he were very hungry; we can scarcely conceive of his eating a third loaf. Hence the marginal utility would become zero with the third loaf. But this is obviously not in harmony with everyday facts. If a man should find that the price of the ordinary loaf of bread was temporarily one cent, he would not refuse to pay that price for a third loaf because he could not eat it at once; he would lay in a considerable stock for his future needs, showing that the marginal utility when he had three loaves was not zero but considerably more than one cent. Ultimately he would have so many loaves that he would not pay one cent for another. This would be when he had more than he could easily carry home or store, or more than would keep. The marginal utility would thus get down to one cent and would finally reach zero, when he would not take another loaf as a gift. But this point is not reached till long after the point at which immediate hunger is satisfied. To the psychological law of satiety must therefore be added the principle that man takes account of the future and makes provision for his needs beyond the present moment.

Marginal utility and total utility. Returning to the analysis of utility, we recall that the marginal utility is the valuation placed upon one of the units on hand. The units are all alike; hence the estimate placed on any one is the marginal utility. But it does not follow that the whole stock would be valued at that rate. Having parted with one unit, the quantity on hand is reduced and there is a new marginal utility higher than before. Consider the curve *am* again (Figure 30). Suppose the stock on hand consists of five units. The marginal utility, the price at which any one unit would be surrendered, is two dollars and a half, measured by the line *se*. But the whole stock of five would certainly not be given up for five times two dollars and a half. Having parted with one, there would be only four on hand, and the marginal utility would be measured by *rd*; that is, four dollars. The total utility is therefore not the product of the marginal utility by the number of units, but it is rather the sum of the several marginal utilities for each quantity from one to the number actually in hand. The total utility of these five units would therefore be measured by the sum of the lines *na*, *pb*, *qc*, *rd*, and *se*. It would be $10 + 8 + 6 + 4 + 2.5$, or 30.5; that is, $30.50. This is reasonable, since total utility should obviously not be measured solely by the least important want satisfied. It should represent the sum of the valuations placed upon all the wants satisfied.

Some apparent price anomalies resolved. Understanding the relation between marginal utility and total utility, we are able to explain certain price relations which sometimes appear as puzzles. A diamond may be worth a thousand dollars; a loaf of bread is worth ten cents. How is it that some of the most useful things in the world are so cheap, while other things that could be dispensed with without any real suffering are so expensive? Why is a prime necessity of life worth less than a mere luxury? To some this seems strange, inexplicable, or even wrong. The answer is that the value of anything depends, not upon somebody's judgment of the importance or righteousness of the want satisfied, but on demand and supply, and on the demand side it is marginal utility that controls. Marginal utility, we have just seen, varies inversely with the quantity. Diamonds are scarce; their marginal utility is therefore high. Bread is very abundant; its marginal utility is correspondingly low. People are willing therefore to pay high prices for diamonds, but very little for a loaf of bread. Of course the total utility of all the bread in the world is very great, even though the marginal utility is small.

Abundance and scarcity. Some of the most useful things in the world are so abundant that they have no marginal utility at all and therefore no value. Among the prime necessities which man must have in order to sustain life is air; nothing could be more useful. Yet no one will pay for air, simply because it is so abundant that everyone can have all he wants for nothing. Its marginal utility is zero. To have value, a thing must have not only utility, but a marginal utility greater than zero. The only things that can have marginal utility are those which both have utility and are so limited in quantity that there is not enough to satisfy everybody's wants. It is for this reason that some of the most important things in the world, as sunlight, air, sometimes water, are not included in the subject matter of economics; they are not economic goods. They have no value, are not bought and sold. It is this, rather than any permanent physical characteristic, that rules them out. Water, free in abundance for all in many country districts, is limited in the city. It there has a marginal utility, has value, and is an economic good. Land in certain newly settled regions is sometimes so abundant as to be valueless. This condition is almost sure to pass, and in general land is so strictly limited and so insufficient to satisfy all human wants that its marginal utility and its value are usually high.

Individual demand. From marginal utility it is an easy step to individual demand. By individual demand we mean of course a schedule of the respective quantities of a good which an individual would choose to take at all possible prices. The curve *am*, in Figure 31, is the curve of marginal utility, identical with the curve in Figure 30. It shows the estimation of the loss which a certain individual would suffer by parting with one golf ball, assuming him to have any given number on hand. But it can just as well show the price he would be willing to pay for one additional golf ball. For example if, having three units, the utility of one golf

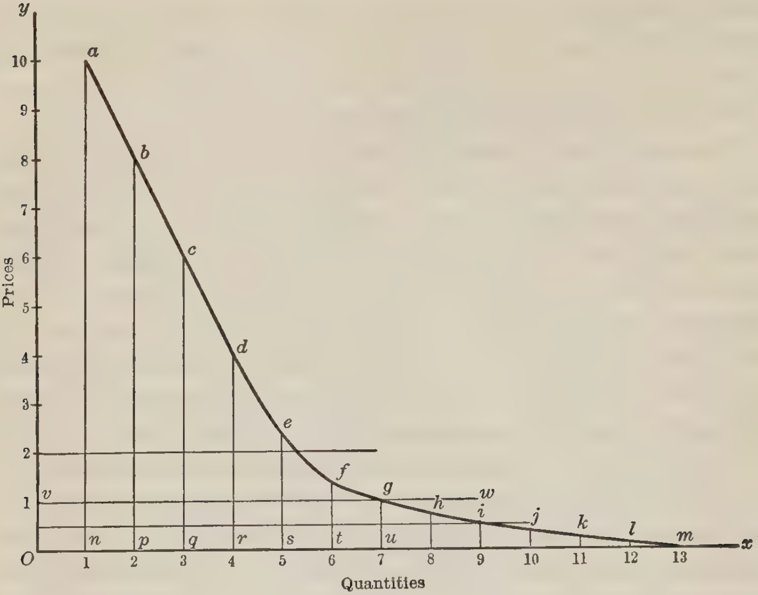


FIG. 31.

ball is six dollars (*qc*), then, having two units, one would be willing to pay six dollars for a third. Every ordinate thus measures the price one would be willing to pay for the additional unit which would bring his stock up to the number indicated on the base line, provided of course he were held up each time for his highest price.

Now that of course is not the way golf balls, or other things, are ordinarily bought. There is an established price on the market. No matter how badly the buyer may want the first few units, no matter how great their utility to him, no matter how much he might be willing to give for them if he had to, he does not have to pay more than the market price. And he can ordinarily buy all he wants to at that price. The question he has to determine is, not what price he will pay, but how many he will take at the established price. The decision is determined by the marginal utility curve. Suppose that our sportsman, laying in a stock of golf balls for the week end, finds the price of the particular make he uses is one dollar. How many will he buy? He will certainly buy one; rather than go without one he would be willing to pay ten dollars, which would measure the marginal utility if he had one only. Similarly, since the utility of a second golf ball is eight dollars, he would certainly buy the second. Likewise he would buy a third, a fourth, a fifth, and a sixth, since in each case the utility to him is greater than the price. The utility of a seventh golf ball is just one dollar. He will just get "his money's worth" by buying this one. The utility of an eighth ball is seventy-five cents. Since it would clearly be foolish to pay a dollar for what has only seventy-five cents' worth of utility to him, he will not buy an eighth. He would gladly have this one and several more, but not at a dollar. We conclude that this person would buy at least six golf balls; he would not buy eight. It is a matter of indifference whether he buys the seventh or not.

Graphic illustration. All this is clearly shown by the curve of marginal utility. From the point on the y axis which measures one dollar, we draw a horizontal line, vw. This line intersects the curve at the point g, indicating that at a price of one dollar seven golf balls will be bought. Again suppose our golf player had found that they were charging two dollars. He would consider this an exorbitant price and, while he would have to get enough to carry him through the week end, he would buy fewer than if the price were more reasonable. A glance at the diagram shows that he would buy five. If he bought six or more, the marginal utility would be less than the price, whereas the fifth has a utility somewhat greater than the price. If on the other hand he had found that his favorite brand was being sold at fifty cents, a bargain price, he would have bought a few more. According to the diagram he would have taken nine at that price. Given an individual's curve of marginal utility for any particular good, we can tell how many units he would buy at a given price. The curve of marginal utility is thus also the individual demand curve.

Obtaining the maximum total utility. Recalling what we have learned of marginal utility and total utility, it is evident that each unit of any particular good purchased up to the last one gives the purchaser a greater utility than the price he pays. On the last one he "breaks even," or as near to that as the size of the unit will permit. There is thus a gain upon each unit (except possibly the last) which is called the *buyer's surplus.* As he buys more units, this unit gain becomes steadily less, but the diminishing gain on the later units detracts nothing from that already obtained on the earlier units, and the total utility and the total surplus increase till they reach their maxima when the last unit is bought, having a utility equal to the price paid. The diagram in Figure 31 will help us again. If seven golf balls are bought at one dollar apiece, it is because the marginal utility to the purchaser is one dollar when he has seven. The various ordinates, as we have seen, measure respectively the utility to him of one golf ball if he had only one, of the second one if he had only two, and so on; in other words, the respective maximum amounts he would have consented to pay if he had purchased his seven golf balls one at a time and been held up each time. The line *vw* cuts off from each of these amounts the amount actually paid for that unit, one dollar; therefore the part of each ordinate above the line *vw* measures the buyer's surplus on that unit. This unit surplus diminishes with each successive unit, but does not become zero till the last one bought, and if the purchaser should stop buying before this he would clearly sacrifice somewhat of the total utility that is within his reach.

How both parties to a trade can gain. It is this principle which explains, what has sometimes been difficult for people to believe, that both parties to a trade can gain. It has been somewhat plausibly argued that, since in any honest trade the values of the two things exchanged must be the same, neither party can have made a gain; one party can gain only by giving less value than he receives; that is, by cheating the other party. It was partly this notion which led the churchmen and lawmakers of the Middle Ages to believe that trade should be strictly regulated and customary prices enforced in order that neither party to trade should gain at the expense of the other.

We can see now that what the purchaser gets from a trade is not measured by the value received; *i.e.*, the price multiplied by the number of units, but is the total utility received, which is always greater than the value or the money paid. The seller on the other hand, has in the money received a utility at least equal to and generally greater than the utility of the goods with which he parted, though this side of the trade will be more clearly set forth in the next chapter, of which it is the subject.

The decision to buy. We thus arrive at the conclusion that in general a person will buy any good as long as its marginal utility to him is equal to or greater than the utility of what he has to pay for one unit of it; *i.e.*, the price. He will stop when the marginal utility is less than the utility of the price. The decision as to any purchase involves the comparison of two separate utilities, the marginal utility of the thing to be bought and the utility of the price; that is, of the money that must be paid. Choice must be made between these two utilities. This means choice between the thing to be bought and the most desirable of all the other things that could be bought with that amount of money. A man is considering whether to spend a couple of dollars for a ticket to the matinée. He thinks that he might instead go to the baseball game; he also needs a new straw hat, and it would be a pleasure to take a two-pound box of chocolates home to the family; and there is still another possible use of the two dollars; *i.e.*, depositing it in the savings bank — saving it. The question whether to go to the theatre is a question of the marginal utilities of all these several competing uses for two dollars. A rational choice will devote the money to that use whose marginal utility is the greatest. The marginal utility curves of all the things that one might buy thus enter into the problem of any particular purchase. If for example it is finally decided that more satisfaction would be derived from the purchase of a straw hat than from any other possible use of two dollars, it follows that the marginal utility of theatre tickets is less than the marginal utility of straw hats, and the ticket is not bought. In this way all possible uses of one's money are set off against each other. By spending each dollar for the thing that will give him the greatest utility, one distributes his money among the countless rival wants calling for satisfaction, and, since the quantity purchased of each good is subject to the test of maximum utility from that good, it is evident that rational spending enables each individual to obtain the maximum total utility — the maximum satisfaction of his wants — which his means will permit.

Back to total demand. We thus arrive at the principles which govern the purchases of an individual, which determine how much he will pay for anything and how much of anything he will buy; that is, the principles of individual demand. The actual demand for any good in any market is the sum of the demands of all individuals who are contemplating buying that commodity in that market. The quantity that would be taken at each price is the sum of the respective quantities that each of the individuals would take at that price. So the demand schedule and the demand curve are the sums of many little individual schedules and curves.

Real people. It is hardly necessary to remind the reader that the ordinary individual does not sit down and carefully plot his marginal utility curve before he decides how many tons of coal or pairs of shoes or pounds of sugar he will buy. Purchasers do not usually base their judgments on any such precise mathematical calculations as have been assumed in our illustrations. Poor people and even the moderately well-to-do often appear to spend their limited stores of money in the most foolish and irrational manner. A family that is without means to provide warm clothing for the coming winter will cheerfully spend ten dollars on a day's outing at the seashore. The victrola, with its monthly installment payments, is eating up the money that would seem much better spent for nourishing food. The daily attendance on the "movies" seems quite superior to any law of diminishing utility. Clever advertising and "smart" salesmanship bank on the careless and irrational way in which people spend their money. All sorts of appeals are used to induce the careless purchaser to buy things whose utility seems absurdly small in comparison with the money expended. To some the appeal of a "bargain" is simply irresistible, and people come home laden with things for which they have no real need, bought simply because they were cheap; every retail merchant takes advantage of this human characteristic. In any city it is possible to see people buying things in one store when the identical articles are on sale at a lower price only a block or two away. A study of the purchasing habits of the general public is apt to leave one with a low regard for average human intelligence.

Is all buying irrational? From such facts as these, which might be multiplied indefinitely, the conclusion is sometimes drawn that there is no such thing as rational judgment in buying, that the law of marginal utility has no existence outside the imagination of the theoretical economist, and that all our talk of demand and supply is pure theory without any useful application to the hard facts of the practical world.

But let us not be too hasty in jumping at such conclusions. For one thing, it will be well to recognize that much of the conduct of other people which we regard as irrational is not really so but is merely at variance with our own judgment. We readily think we could lay out a family budget for our impecunious neighbor which would materially increase the utility of his purchases. But it is his judgment (such as it is), not ours, which decides how he can get the most for his money. The well-fed, well-dressed woman of the "upper class" points with scorn at the "irrational" conduct of the stenographer taking her pathetically inadequate luncheon at the soda fountain in order to save money to buy silk stockings. But this conduct, whatever else may be said of it, is not irrational. It is the height of careful judgment. The girl wants the silk stockings more than she wants a nourishing lunch; she has planned carefully how she can save the price; she is acting strictly in accordance with the law of marginal utility. Her conduct may be "foolish" according to some other person's judgment. But the economist is investigating how she actually spends her money, not in how she "ought" to spend it. Let us reverse the case. The shopgirl laughs at the wealthy patrons of the Fifth Avenue store who "haven't sense enough" to walk over to Sixth Avenue and buy the same things at lower prices. But the wealthy shopper is not irrational. She wants the calm restfulness, the air of luxury and dignity, the deferential personal attention of the exclusive Fifth Avenue shop. These are a part of what she is paying for.

She knows she could save money on Sixth Avenue, but the loss of these other advantages would mean a net sacrifice in the utility of her purchases. She also may be "foolish," but that is beside the present point.

Economic laws: the margin of error. So it is with much of the conduct of the general public. Acts which to a superficial glance appear irrational turn out to be based upon real judgment, however "foolish" or "wrong." True, there is conduct which is undoubtedly irrational, as has been pointed out. But this is not true of the majority of human acts, and in large measure the acts of the irrational individuals tend to offset each other when merged in the composite action of the mass. The irrational exceptions are not sufficient to destroy the validity or practical importance of our economic generalizations or laws. What we finally get is simply another example of the margin of error in their practical application. In this respect, as we have before had occasion to observe, economics differs from the "more exact" sciences only in degree. The physicist can explain or even predict the course of a projectile by reference to his laws of motion, gravitation, air resistance, and so on. But there will always be a certain margin of error in the actual result. So our economic generalizations, such as the laws of demand and supply and price, are useful for the explanation and even the prediction of practical matters, in spite of a margin of error which is sometimes small, sometimes great. Every wise business man makes it his task to study his clientele. He finds that people do actually behave in certain ways. They react to certain stimuli in ways that may be predicted with only a moderate margin of error. These are hard practical truths, which the business man neglects at his peril.

The law of demand. Among these practical truths we find the economic laws of marginal utility, demand and supply, and price. People as a rule actually do determine their purchases in accordance with the law of marginal utility. They will ordinarily buy more if the price is low, less, if it is high. They have some idea of when to stop buying because more enjoyment may be obtained from other things which their money will buy. They realize that there are some things that they cannot afford at present prices. They alter their purchases with price changes, buying the fruits and vegetables which are cheap because in season, and turning to other forms of food when these prices rise. Nobody buys with perfect judgment at all times. Some are more rational than others. A few may appear quite irrational. But the "freakish" purchaser is the exception, much irrational conduct is mutually offsetting, and human conduct as a whole is rational and subject to reliable generalizations or laws. So, while our generalizations as to marginal utility and individual demand are only approximations, not holding true of every individual or of every purchase of any individual, they still hold true in the main, and from them we can derive a law of (total) demand which is both true and practically important.

Demand is responsive to countless influences. Much more might be written about demand. The circumstances which determine people's purchases: custom, tradition, education, fashion, individual peculiarities, sudden whims, selfishness, altruism, patriotism, religion — all this furnishes an interesting field of study. Equally interesting is the subject of the creation or modification of demand, by sellers through advertising, salesmanship, etc., or by "reformers" through education, religion, and public appeal of various sorts, or by laws forbidding some purchases (as alcoholic liquors) and compelling others (as automobile license plates and bathing suits). On the foundation of the brief analysis of human wants already presented could be built an extremely interesting study of how each want works itself out, in correspondence and in competition with other wants, in the demand for particular goods. With the introduction now before him, the reader can be left to follow out such lines of study for himself.

CHAPTER XVII

COST OF PRODUCTION, SUPPLY, AND COMPETITIVE PRICE

We have observed that the immediate factors which determine competitive prices are demand and supply, cost of production, and now and then the decision of a sole buyer. Back of these immediate agents, a host of circumstances play their part in the determination of price. In the chapter just preceding we have examined some of these forces on the buyer's side of the market. We now turn to the seller's side, in order to consider the more important influences back of cost of production and supply.

The different viewpoints of buyers and sellers. The buyers of goods are either the consumers, who seek goods to satisfy their wants, or the producers,[1] who buy in order sooner or later to sell either to the consumers or to other producers. In any case the goods finally reach the consumers, and it is this last exchange which is the fundamental one for the study of the demand side of the forces determining price. The demand of manufacturers and dealers is not the ultimate demand; it is an indirect or reflected demand, arising from the demand of the "ultimate consumers." In seeking the causes back of demand therefore we rightly passed over the demand of the producers in order to give our attention to the causes which determine the demand of the ultimate consumers; that is, of those who seek goods to satisfy their own wants. When now we turn to the study of supply, we meet at the outset the important fact that in general there is no corresponding class of sellers who are parting with goods which might satisfy their own wants. Practically all selling is by producers, not by users.

Of course there are exceptions. Financial reverses or other changes in personal or family circumstances sometimes lead people to sell their personal effects, thus depriving themselves of their use in order to devote the money received to the satisfaction of more important wants. Quite similar is the case of goods which are not capable of reproduction, such as rare postage stamps, first editions of books, antique furniture, paintings by old masters, and all articles which are desired on account of some peculiar association or quality which cannot be reproduced. When any such thing is offered for sale, someone must have been prepared to give up its use in order that the purchaser may enjoy it. The supply in these cases is, like demand, largely based on marginal utility — *i.e.*, to the sellers — coupled with the sellers' judgment as to what prices buyers are likely to offer. This will be sufficient explanation of the supply of non-reproducible goods and goods offered for sale for personal reasons by the users. These are obviously exceptional cases, not representative of the ordinary run of sales. Our attention for the remainder of this chapter will be devoted to those goods which are being currently reproduced or are at least capable of reproduction.

---

[1] The term producers includes of course merchants and all others who have a part in the productive process.

For goods in general there is thus an important difference between the ultimate buyers and the sellers. The purchaser buys a set of golf clubs in order to play the game. The merchant who sells him the clubs is not thereby giving up his opportunity to play golf, and the estimate he places upon the clubs, unlike that of the buyer, has nothing to do with the pleasure of playing golf. The principle of marginal utility is not the guide to the seller's conduct.[1]

The retail dealer, from whom the consumer buys, is in the business of selling goods for a profit. His profit depends on the one hand upon the prices at which he is able to sell his goods and on the other hand upon the prices he has to pay for goods, for wages, for interest, and for all the other items in the cost of conducting his business. These are the facts that determine his conduct in relation to price, and the same is true of the wholesaler, the jobber, the manufacturer, the farmer, and all other persons engaged in the various steps of production from the extraction of the raw material to the delivery of the finished good to the final consumer. Each of these is a seller of goods or services, and it is from them that the supply comes.

---

[1] It is true that we must finally push our analysis of the selling side of the market back through the actual sellers to the "ultimate producers," corresponding to the ultimate consumers on the buying side. We thus reduce all costs to labor cost and the cost of waiting for future income, and we come finally to a principle of "marginal dis-utility of labor," somewhat analogous at least to the principle of marginal utility of goods. This analysis we shall undertake later, when we come to the study of wages, profits, and interest. It must wait however till we have disposed of our present problem, which has to do with the forces which immediately control the actual sellers of goods.

Business costs and the lower limit to price. Since business men are engaged in production for the sake of profits, they will not, if they can help it, offer their products for sale except at prices at least high enough to cover all their costs. The producer who is unable to sell his product at prices high enough in the long run to cover his costs will sooner or later give up and retire from the business. And it must be recognized that among the costs of business must be included a reward for the entrepreneur's own labor of management. This is clear enough when the manager is a hired employee of the owner of the business who himself takes no active part in its conduct. But it is no less true of the entrepreneur who manages his own business. The entrepreneur generally has before him the two alternatives of continuing to run his own business or of hiring out to an employer. His choice will ordinarily depend upon the relation between the wages or salary which he could obtain as a hired worker and the income obtainable from his business. Unless the latter is in the long run at least as great as the wages or salary he could command, the business man will ordinarily not continue as an independent producer. Production cannot be carried on under our present economic system without the services of the business manager. The minimum reward required to induce him to perform his function is therefore just as much a necessary cost of production as the wages of laborers or any other cost. This minimum is sometimes called "wages of management," to indicate that this part of the income of a business man is only what he could presumably earn as salary or wages if he should choose to work for an employer instead of running his own business. Whatever gain he may receive above this minimum is called "profits." Hereafter in speaking of costs of production it will be understood that we include "wages of management," whether received by paid managers or independent entrepreneurs.

The upper limit to price under competition. The costs of production thus set a price minimum below which producers will not in the long run offer to sell their products. Of course producers would gladly receive and will generally seek to obtain prices higher than this minimum. Are they restrained by a corresponding upper limit to price? The answer will depend upon whether production is carried on under conditions of competition or of monopoly. In the present chapter we shall devote our attention to competitive production. There is competitive production of a commodity or service when two or more independent producers are engaged in its production in competition with one another.

Let us assume that several competing bakeries are offering bread to the consumers of a certain community. For simplicity we may assume that the size of loaf and the quality of the bread are practically uniform, and we may further assume that the cost of production is practically the same for all the bakeries. This, according to our present analysis, sets the limit below which bread will not in the long run be sold. What we are now seeking is the upper limit to the price of bread. Each baker is striving to make his profits as large as possible. Large profits would result either from a high price or from large sales. The effect of charging a high price however is to drive customers away to rival bakers whose prices are not so high, whereas the one best way to increase sales is to offer lower prices than are offered by one's rivals. Each competing producer thus finds himself impelled, in the struggle to hold and increase his trade, to make his price as low as possible while still covering all his costs. Any producer who persisted in offering goods only at prices materially above the cost of production would find his customers leaving him and would ultimately be forced out of business.

Price determined by cost. Under the spur of competition therefore cost of production sets not only the minimum limit but also the maximum limit to the prices at which producers will in the long run offer their products for sale, and we arrive at the generalization that producers will offer goods and services at prices which in the long run will tend to be equal to the cost of production. In the development of this conclusion we have had before us principally the manufacturer. But it will be recognized that our conclusions will be equally applicable to the farmer, the miner, the wholesale dealer, the retailer, and to every person engaged in the long and intricate process of production from the extraction of the raw material till the finished product is delivered to the final consumer. Each entrepreneur must in the long run charge prices high enough to cover his costs, including his minimum wages of management; and no entrepreneur in the whole series can in the long run demand prices materially above the costs of production, always assuming that practically unrestricted competition prevails.

Prices not fixed by cost of existing goods. But the cost of production to which reference is made in the preceding paragraphs is not past cost already incurred. It is evident that, at any given time, a product will not necessarily be offered for sale for what it cost to produce it. This is obviously true of the whole group of non-reproducible goods. A rare postage stamp sells without any reference to its cost of production, as does also a painting of a famous old master or a rare and beautiful piece of antique furniture. Such things will be offered only at prices far above their cost. On the other hand a painting by a past nonentity may be offered for a good deal less than it cost to produce it.

The proposition is also true of reproducible goods. Suppose a new invention has greatly reduced the cost of making a certain product. Producers and dealers will generally find it impossible to dispose of all their goods on hand at a price equal to what it cost to produce the existing stocks. The fact that the stock may now be replenished at a lower cost will lead competing producers to offer the product at a price related to the new cost of production. Those who are stocked up with the product will find themselves compelled to offer their stocks at the new price. On the other hand consider the case of a change which increases the cost of a certain commodity, such as a rise in freight rates or an agreement fixing a higher scale of wages. There will now be a new and higher cost scale governing all future production. Those who have stocks of the commodity on hand, knowing that they cannot be replenished except at greater cost, will not be likely to go on offering the commodity for sale at a price just covering the old cost of production. They will find that they can get a price governed by the new cost even for those articles which were produced at the former lower cost, and this will control the supply offered by them.

Future cost controls. Producers, in deciding the prices at which they will offer goods, are therefore controlled by future costs of reproduction and very little, if at all, by the past costs at which existing stocks were produced. Of course no wise producer closes his eyes to the past or present scale of costs, and if there appears no reason to anticipate a change in the near future, supply is apparently determined on the basis of past costs. But that is simply because costs already experienced have turned out to be the same as the later sales price, and it is only as an indication of what future costs may be expected to be that past or present costs enter into the problem. In making offers for future delivery, producers are governed by their estimates of future conditions of production and future costs, and the selling prices at which they thus arrive will generally determine also the prices at which existing stocks will be offered.

In spite of the logic by which this conclusion was reached, the reader may possibly find it hard to give up the notion that manufacturers and dealers, through their modern systems of cost accounting, do fix the prices of their products on the basis of what it cost to produce them plus a pre-determined rate of profit. Cost accounting does it is true have its effect, particularly when, cost conditions being for the time comparatively stable, there appears no reason for expecting that future costs will be materially different from those of the immediate past. But the final authority of future costs is not overthrown, as a careful analysis of the " load " added for profits will often disclose.

No system of cost accounting will long prevent producers from taking advantage of changed conditions which enable them to make more than the usual or expected profit on existing stocks, as they did for example in the boom period of rising prices that culminated in 1920. And by the same token producers cannot, by the simple device of fixing their selling prices on the basis of "cost plus," avoid selling existing stocks at a loss when changed conditions lower the scale of future costs. On one occasion a reduction of the wholesale price of gasoline led the retail sellers in New Haven to lower their price from twenty-eight to twenty-six cents. One dealer however persisted in holding to the former price. When questioned about it, he replied that he could not without loss reduce his selling price till he had disposed of his existing stock purchased at the higher wholesale price. Of course the result was to drive trade to his competitors and to cause him a greater loss on new business sacrificed than he would have incurred by selling his existing stock at the new price regardless of its past cost.

Conclusion and qualifications. We thus reach the conclusion that the supply of producible goods is governed by cost of production, meaning expected future cost rather than past cost. Of course there are exceptions. A change in fashion may leave a stock of some article for which there is now little use and of which there will be no further production. A search through the barns and carriage houses of the United States would disclose many an old-time sleigh. Sleighing has pretty much gone out of fashion, and the fact that a sleigh, still in good condition, might now cost three hundred dollars to reproduce will not enable its owner to get any such price for it. He would gladly offer it for sale at a price far below its cost. Dealers sometimes find themselves with unsold stocks which cannot be held over and which they offer at "sacrifice" prices having no relation either to past or future costs of production. A panic or industrial crisis may cause sales at less than cost of reproduction. On the other hand, an unexpected demand for some commodity by persons who will not wait for future production may enable the fortunate dealers to offer their limited stocks at exorbitant prices without regard to cost. Numerous such examples will readily occur to the reader. But these are the exceptional cases, and they do not diminish the significance of the general principle that in the long run future cost of production, as visualized by the entrepreneurs, controls supply.

tother

Analysis of cost. Our next task is to obtain a definite idea of the meaning of unit cost of production, in particular to answer the question whether the cost is the same for all the units turned out by any given producer. For this we shall need a careful analysis of cost. Let us suppose that a certain manufacturer is producing shoes, using a plant (land, factory, machinery, etc.), which cost $1,000,000. The interest on this capital is of course one of the elements in the cost of producing shoes. It is obvious that, this sum having once been invested in the plant, the interest will go on indefinitely, without regard to the number of shoes produced. Indeed the interest will continue even though no shoes at all are produced; it cannot be escaped even by closing down the factory. While this particular owner might perhaps escape by selling the plant, the purchaser would have to assume the burden. There is another group of costs whose amount has little or no dependence upon the number of shoes produced; this group includes such costs as the salaries of the principal officers and department managers of the factory, the payments for insurance, for lighting and heating the factory, etc. These costs are unlike the interest on capital however in that they would generally cease if the plant were closed down. Finally there are costs which are related closely to the quantity of the product, so that their amount varies almost if not quite in proportion to the amount produced. Thus the cost of leather and other materials, wages, and certain other costs will vary in amount according to the number of shoes produced by the factory.

We therefore divide production costs into three classes: (1) *fixed costs,* meaning those costs which must be borne indefinitely, regardless of the amount of the product and even though the plant should be shut down, (2) *general running costs,* or *overhead costs,* which would not have to be incurred if the business were given up, but whose amount has little or no relation to the amount of the output, and (3) *particular running costs,* whose amount varies, approximately at least, in proportion to the amount of goods produced.

Relation between each producer's output and unit cost. A numerical example will help us to arrive at the relation between these several groups of costs and the total cost of production per unit of product. Returning to our example of a shoe factory with $1,000,000 of capital invested in plant, let us assume that the annual interest charge, at 6 per cent, is $60,000. Let us assume further that the overhead costs are $450,000 a year and that the particular running costs are $3 per pair of shoes produced. What will be the total cost of making a pair of shoes? Evidently it will vary according to the number of pairs produced. The following table shows the results of three separate assumptions as to the output:

COST OF PRODUCING A PAIR OF SHOES

	First Assumption	*Second Assumption*	*Third Assumption*
Amount of product	500,000 prs.	1,000,000 prs.	2,000,000 prs.
Fixed costs	$ 60,000	$ 60,000	$ 60,000
General running costs (overhead)	450,000	450,000	450,000
Particular running costs ($3 a pair)	1,500,000	3,000,000	6,000,000
Total costs	$2,010,000	$3,510,000	$6,510,000
Costs per pair	$4.02	$3.51	$3.26

This illustration brings out with clearness the principle that, under such circumstances, the average unit cost of production decreases with an increase in the amount produced. The reason is simple enough; *i.e.*, since the fixed costs and overhead costs are constant in amount, the part of such costs assigned to each unit becomes successively less as the number of units increases. There is therefore a general tendency for the unit cost of production of each producer to decrease as he is able to increase his product.

The limit to reduction of unit cost. But there is a very definite limit to this tendency. In our illustration one million pairs of shoes could be produced more cheaply per pair than half a million, and two million pairs more cheaply than one million. But it does not follow that there would be further reduction for three, four, ten, twenty million pairs. Evidently when half a million pairs were made, the plant and the services for which overhead costs were incurred were not being fully utilized; otherwise it would not have been possible to increase the output to a million pairs and then to two millions with the same plant and the same general services. Sooner or later, if production is continuously increased, the time must come when plant and general services are fully utilized. Further increase in the product will require additional investment in the enterprise, either by enlarging the existing plant or adding a new plant, and will also require an increase in the staff of executive officers, etc. This will mean an increase in the fixed and overhead costs, and so in the total cost per unit. When the capital investment and the executive staff are increased, the increase will, at the start, almost certainly be greater than sufficient to care for the increased product which can be immediately sold. Therefore there will again be a situation of unutilized plant, with the possibility of another succession of reductions in unit cost as output is further increased. This tendency may be illustrated by a curve, as in Figure 32. The ordinates, such as *ab*, measure the average cost of producing one unit when the number of units produced is measured by the corresponding abscissas, as *cb*. When the quantity reaches *od*, the enterprise is fully utilized, and, immediately after the plant has been enlarged or another plant added, the unit cost shoots up as measured by *df*, followed by another period of decline, and so on with each increase in the enterprise.

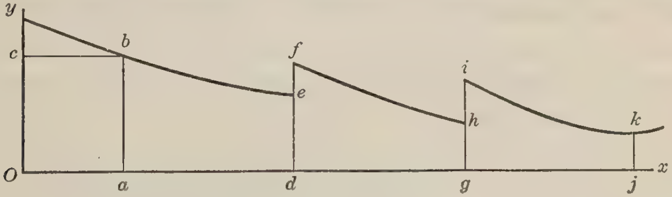


FIG. 32.

Now the reader will recall that there is in many lines of industry a gain in efficiency from large scale production. It is likely therefore that after the enlargement of the enterprise the unit cost of production will decline to a lower point than was reached when the plant was previously fully utilized. Referring to the diagram, *gh* is likely to be less than *de*. But this will not go on indefinitely, since there is a limit to the size of the business unit, beyond which efficiency declines. These are applications of principles investigated in our earlier study of the subject of large scale production (see Chapter VII). Ultimately, after a series of additions, the enterprise grows up to the size of maximum efficiency, after which there will be no further increase in size, assuming general industrial conditions to remain unchanged. For example we may assume that a business unit sufficient to produce the quantity *oj* is the maximum size for efficiency. When this plant is fully utilized there will, other things being equal, be no further attempt to increase the product. The entrepreneur settles down to a product of *oj* at a unit cost of production *jk*.

When all the producers in any industry are in this state; *i.e.* with enterprises of the size of maximum efficiency and plants fully utilized, a limit has been reached, beyond which further increase of production can come only from (1) an increased output of the existing business units, necessarily at greater unit cost (note that the curve in Figure 32 ascends after the point *k*), or (2) the starting up of new enterprises.

When the limit is reached. The principles thus far developed from our analysis of production costs are general, holding good of every kind of industry. When however we inquire how soon the limit described in the preceding paragraph is reached and what follows it, we shall find fundamentally different results of great practical import, depending upon the type of industry concerned.

In certain lines of industry it appears that the business unit of maximum efficiency is an enormous organization, and we have a few huge concerns, " trusts," combinations, etc., each with a very great output. Under such conditions frequent additions to the plant of any particular concern may be possible, additions which may involve the investment of millions or hundreds of millions of dollars of capital. And again, under such circumstances, the existence of large business units not fully utilized may be quite common and may long persist. Thus there is opportunity for producers to decrease their unit costs if they can increase their product, and the industry as a whole has not yet reached the point where decision must be made between increasing the product of existing concerns at a higher unit cost or starting new enterprises.

But of industry in general the foregoing account does not appear to hold. In agriculture, to start with, the business unit of maximum efficiency is quite small, and we generally find production conducted on thousands of small farms, rather than on a few huge estates. Moreover, largely on account of the great number and small size of farms, the plant of each farm is normally quite fully utilized. The period during which an increase in the output of a given farm may be accompanied by a reduction in unit cost is generally soon passed. And the whole industry long ago reached the limit at which further production had to come in part at least from bringing new farms under cultivation. The same is true of many, probably the great majority of, industries in the field of manufacturing, trading, etc. A great number of small or moderate sized business units is the normal condition, and each may be presumed to be fully utilized. Choice has long ago had to be made between increasing the product of existing enterprises or bringing in new ones.

When the limit has been reached (that is, when all business units are of the most efficient size and are fully engaged) and there comes a call for additional production, each producer recognizes that further enlargement of his enterprise would be uneconomical and also that the attempt to increase the product of his existing enterprise would succeed only at the expense of an increase in unit cost. Even so manufacturers, for example, are sometimes induced to increase their output. This may be accomplished for a while by changes in organization, speeding up here and there, calling upon the workmen to give overtime, and so on. Under special stimulus, as for example the urgent demands of the government in time of war, productive concerns may thus be "speeded up" to a considerable increase of product. But this is accomplished only by adding to the unit cost of production, and in most industries it is seldom worth while to go far in this direction. When the existing concerns have grown to the size of maximum efficiency and are all fully utilized, it is preferable, because cheaper, to obtain additional product by starting up entirely new enterprises rather than by pushing the existing ones much beyond the point of greatest efficiency.

To this rule there is one great exception. In agriculture and the other extractive industries land plays the predominant rôle as it does in no other field of production. When the farms of a community are of the size of greatest efficiency and have their plants fully utilized (which we have seen is the normal condition) it is not such a simple matter to obtain additional product by starting up new farms. There is only a very limited amount of idle land available, and this land is of inferior quality to that now used; otherwise it would be already in use. New farms, because of the inferiority of the land, will from the beginning have to be operated at higher cost than prevails on the existing farms. At the same time, increasing the product of the existing farms means putting more capital and labor upon each acre of land, thus altering the proportions in which the three factors of production were combined and increasing both the fixed costs and the running costs per unit of product. Since either device for obtaining an increase in agricultural products thus involves increased unit cost, both will be resorted to, and the increase will have to come partly from the cultivation of additional inferior lands and partly from more intensive cultivation of the lands already in use. This holds true of the extractive industries generally.

Relation between product and unit cost of a whole industry. It is evident from what has gone before that the unit cost of production of any good is dependent upon the quantity that is being produced. This dependence gives rise to three possible relations between unit cost and quantity, which are called respectively "increasing costs," "constant costs," and "decreasing costs." Let us examine each.

Increasing costs. In agriculture, after the early stage of abundant fertile land which has long since been passed in every well-settled community, it is possible to increase production only at the expense of higher costs per unit of product. Inferior lands must be pressed into service, more laborers must be employed per acre, more capital must be employed in the form of buildings, farm machinery, tools, drainage systems, etc., more must be spent on fertilizers, and so on. All this means an increased cost per unit of product. Similarly in mining, the output may be increased only through resorting to inferior mines and through working existing mines more intensively by going to greater depths, taking out lower grade ores, installing more expensive machinery, and in other ways adding to the cost for capital and labor more than in proportion to the gain in output. Similar conditions appear in lumbering, fishing, and all the extractive industries by which the raw products of nature are obtained. All these industries are thus normally operated, except in their remote earliest stages, under what is known as the *law of diminishing returns* or, what is the same thing, the *law of increasing costs.*

Constant costs. Outside the extractive industries production, after the point at which existing business units are of the most efficient size and fully engaged, may be increased by starting up new plants, without the necessity of pushing the existing ones on to produce at higher costs. Since the new plants may be presumed very soon to attain a degree of efficiency equal to those already in the field, production may be increased without any increase in the unit cost. This, other things being equal, gives a condition of constant costs, which in a broad way prevails throughout a great part of the whole field of production outside the extractive industries.

We have however to remember that all industries depend finally for their raw materials upon the extractive industries, which are subject to the law of increasing costs. Coal, iron, wood, leather, cotton, and so on — none of these materials can be commanded indefinitely without paying higher costs, which must inevitably tend to increase the cost of the finished products into which they enter. There must therefore be in every industry a limit beyond which constant costs will change to increasing costs, though the imminence of this change may be more or less remote.

Decreasing costs. Finally there are certain lines of industry which are operating, temporarily at any rate, under conditions of decreasing costs. This will be the case in an industry in which the business units have not yet reached the size of greatest efficiency or become fully utilized. In this situation it is possible to secure increased production by utilizing more completely units which are not being worked to full capacity and by increasing the size of those units which have not yet grown up to the size of maximum efficiency. Instead of paying the penalty of higher costs, increased product may actually be obtained at a reduction in unit cost.

The condition of decreasing costs is not a stable one. Its possibility in any industry is generally due either to enterprises only partially utilized or to enterprises smaller than the most efficient size; that is, to more producers than are necessary to meet the demand for the product. Under such conditions, competition is likely to be especially severe, as will later be explained, and there is a tendency toward elimination of the least efficient producers and combination among others, reducing the number of producing units and increasing their size until the point of maximum efficiency or lowest unit cost has been reached. Thereafter the call for more production will be met by adding new plants presumably as efficient as those already in the field and so bringing on the more stable condition of constant costs, which appears to be the logical end toward which are headed those industries which for the time being are operating under conditions of decreasing costs.

Marginal cost and average cost. In our discussion of unit cost thus far we have had in mind the average cost per unit of product, *i.e.*, *average unit cost.* We need now to introduce another concept; namely, *marginal cost,* by which we mean the additional cost occasioned by producing the last unit of any given number produced. The diagram in Figure 33 will aid the reader to grasp the nature of marginal cost as well as the relation between average cost and marginal cost. The two curves mm' and cc' represent

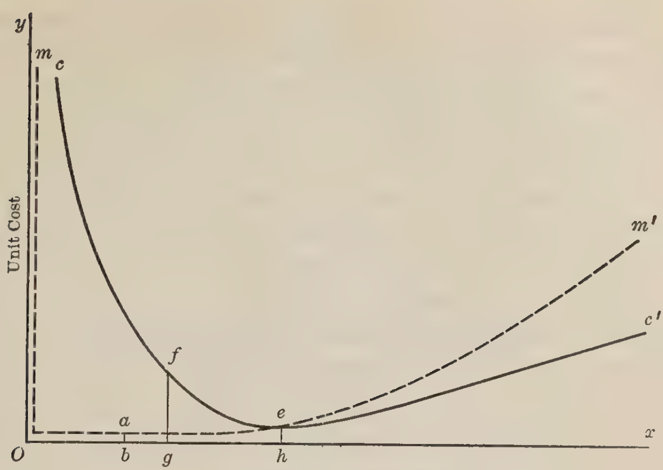


FIG. 33.

respectively marginal cost and average cost for a concern which, for the sake of simplicity, we assume is of the size of maximum efficiency. It should be emphasized that the diagram represents, not a historical sequence, but the present state of affairs, showing the result of producing any given number of units as measured on the base line.[1]

Any point on the marginal cost curve, as *a*, shows by its ordinate, *ab*, the additional cost of producing one unit if the number produced were as measured by the abscissa, *ob*; *ab* then is the marginal cost when the output is *ob*. Within the field of practical assumptions, it is evident that the additional cost of one unit more is practically only the particular running costs per unit, and the marginal cost curve, *mm'*, is therefore a horizontal line until the point where an increase in quantity produced begins to increase the unit running costs, when the curve begins to ascend.

---

[1] Both curves start very high at the left. Indeed it is impossible to draw this part of either curve to scale and keep it within the field of the diagram. If only one unit were produced, that unit would have to bear the entire fixed costs and overhead costs of the concern as well as its unit running cost, which might amount to several thousand dollars. With the second unit produced, marginal cost would drop immediately to practically the unit running cost, while average cost would be about half what it was for one unit. Both curves thus drop rapidly in their early parts and come within the field of the diagram only where a considerable output is being assumed. These initial parts of the curves are obviously of no practical significance.

The curve of average cost, *cc′*, will descend, very rapidly at first, more slowly later, with the increase in the number of units produced, approaching always nearer to the marginal cost curve while the latter is horizontal, and intersecting it soon after it begins to ascend. At the point of intersection, *e*, the average cost is the same as the marginal cost, and thereafter, while both curves ascend, the marginal cost is always greater than the average cost. The point *e* is what we have called the point of complete utilization of plant, when efficiency is at the maximum as indicated by the minimum average cost. Before that point is reached, average unit cost is high because not enough units are being produced to enable the plant to function at its best.

Individual cost curve and supply curve. It is evident that, before his organization is fully occupied, the producer cannot avoid loss unless he receives a price at least equal to his average cost and, since the average cost is higher the fewer units he turns out, he cannot sell at any given price unless he can be sure of producing at least a certain quantity. In order to be able profitably to sell at the price *fg*, the producer must be sure of being able to sell at least *og* units.

After the point *e* has been reached the cost of production rises, and the producer will not be willing to increase his product unless he is offered a price higher than *eh*. Moreover, since the cost of each additional unit will from now on be greater than the average cost, the producer will no longer be willing to increase his product in return for a price only equal to the average cost, since that would mean a loss on the last units produced. In order to find the prices at which various quantities greater than *oh* would be produced, we must follow, beyond the point *e*, the curve *em′* of marginal cost, not the curve of average cost. The marginal cost from this point on is, of course, higher than the average cost. This curve is the normal individual supply curve, showing for any price the quantity that this producer would offer to sell. The curve from *c* to *e* does not show this; it is not a supply curve, but shows rather the miinmum quantity which this producer would have to turn out in order to sell without a loss at any given price.

We note that, as regards the problem of price, the curve of marginal cost has no significance until the point of maximum efficiency is reached. Up to that point it is the average cost curve *ce*, not the curve *me*, which controls. Beyond that point, marginal cost supersedes average cost as the significant factor; it is the curve *em′*, not *ec′*, which controls. In general then the controlling factor is average cost or marginal cost, whichever is the higher. The significant curve is the hybrid average-marginal cost curve, *cem′*.

Supply and price: increasing costs. Now let us suppose that, having the hybrid curve *cem′* for each producer, actual and potential, we add the curves and obtain a similar curve, *css′*, for all producers, as in Figure 34. This is also a hybrid curve. The descending part from *c* to *s* shows for any given price the minimum quantity which must be produced in order that all producers may just break even. It is a curve of average cost. From *s* to *s′*, we have the ordinary supply curve for an industry subject to increasing costs, showing the total quantity that would be offered by all producers at any given price. It is a curve of marginal cost.

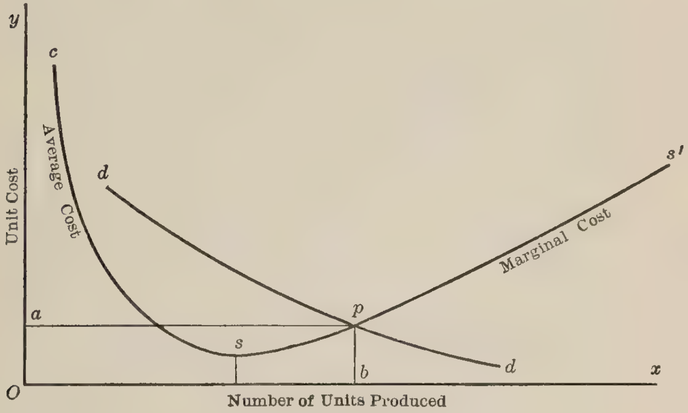


FIG. 34.

It is this curve which, in conjunction with the demand curve *dd*, determines the price, *pb*, and the quantity exchanged, *pa*, as has been fully investigated in earlier chapters. In an industry subject to the law of increasing costs, such as farming, the initial descending curve has no significance. The individual producing units are so many and so small that the minimum quantity which each must turn out in order to sell at a profit is only a drop in the bucket compared with the total quantity which the market demands. A certain farmer may have to raise 5,000 bushels of wheat in order to make his farm pay, but what is that compared with the millions of bushels of wheat which the market is ready to take? Also the industry as a whole has passed the stage, if it ever existed, at which there is any question of a total product sufficient to keep all producers profitably occupied. The demand curve is sure to intersect the other curve well to the right of the point where the ascending supply curve begins.

In any industry subject to the law of increasing costs the individual producer is concerned, not with the minimum which he must sell, but with the maximum which it will pay him to offer. To him the price appears to be the fixed thing, determined by outside causes, and he seeks to govern his production so as to make the greatest possible profit at that price. The wheat farmer for example might produce a small quantity at a cost for each bushel materially less than the selling price. By producing more, the cost of each additional bushel tends to increase. But this will not diminish the profit on the earlier bushels produced at lower cost, and it is to his advantage to keep adding to his output until the cost of the last unit is equal to the expected selling price. This last unit is the marginal unit and its cost the marginal cost. Each producer thus decides the quantities he would be willing to offer at various possible prices, and these individual supplies combine to form the total supply which with demand determines the price and the quantity exchanged.

Thus are explained the forces back of supply in those industries which are subject to the law of increasing costs, among which agriculture and the other extractive industries are predominant, and we conclude that the price tends to equal the marginal anticipated cost of production.

Constant costs. We turn our attention now to those fields of industry in which constant cost is the normal condition. The analysis of individual supply is the same as in the preceding case, and in general the individual cost and supply curves are the same, as illustrated in Figure 33. But when the individual curves are combined, a different supply curve results, because there is not here, as there is in agriculture, anything to prevent a constant influx of new producers as the market demands more product. The existing producers, having each reached the state of maximum efficiency, are not called upon to increase their product at higher unit cost. Rather, the responsibility for the additional product is put upon new producers, who are soon able to produce just as cheaply as those already in the field. Production may thus be pushed very far without any material increase in unit cost, subject only to the universal check of the rising cost of raw materials.

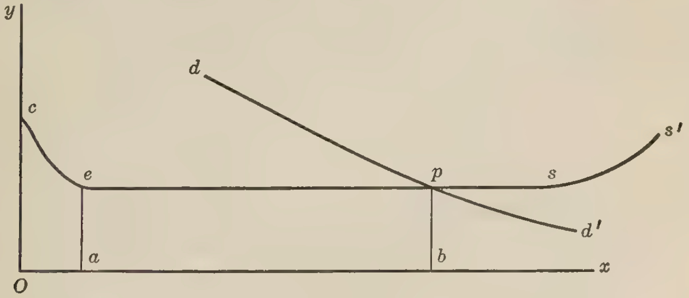


FIG. 35.

The resulting curve is as shown in Figure 35. Like the other, this is a hybrid curve, the initial descending part, *ce*, being a curve of average cost and showing the minimum product necessary to prevent loss to those producers who were in the field when the need of additional product began to bring in new producers. The curve from *e* to *s* is a cost curve indicating that any quantity from *oa* to *ob* could be produced at a cost of *ae*. Since cost is constant and since the producers are so many and generally so small, marginal cost and average cost for this part of the curve are practically identical. The reader will have noted that the curve *cs'* does not become a true supply curve till it begins to ascend at *s*. Before that it fails to do what a supply curve must do; *i.e.*, tell the respective quantities that would be offered by sellers at various prices.

When the demand curve intersects the horizontal curve *es*, we have the case already noted of competitive price determined by cost of production (see Chapter XIII). Demand is not a factor in fixing price, since the price will be the same wherever the demand curve intersects *es* (and of course if the demand curve passes either to the left of *e* or to the right of *s*, the case is no longer one of constant costs). What demand does is to determine, in conjunction with cost, the quantity exchanged. As in the previous case and for the same reasons, the initial descending curve of cost has no significance, either to an individual producer or to the industry as a whole. The demand curve is seen to intersect the supply curve well to the right of the point where cost becomes constant.

The separate producer here, as before, looks upon the prevailing price as a determined thing. Each producer furnishes that amount of the product at which he can operate at maximum efficiency on the basis of that price; he stays in the business as long as he can produce profitably at that price, and new producers enter when they see the opportunity to produce at an average cost not greater than the prevailing price. Thus arises the condition which enables consumers to obtain any desired amount, within limits, without causing an increase in price. This condition generally prevails in the great mass of industries, manufacturing, trading, etc., in which the normal producing unit is comparatively small and where the law of constant costs therefore holds sway. Again we conclude that the price tends to equal the marginal anticipated cost of production.

Decreasing costs. There remains to be considered the somewhat more complex problem of the forces affecting price on the supply side in those industries which are operating under the law of decreasing costs. The individual cost curve is in its general shape like those that have been studied in the two previous cases, except that now the advantages of large scale production are predominant and the producer may be able for a long time to reduce his unit cost in conjunction with increasing product. The

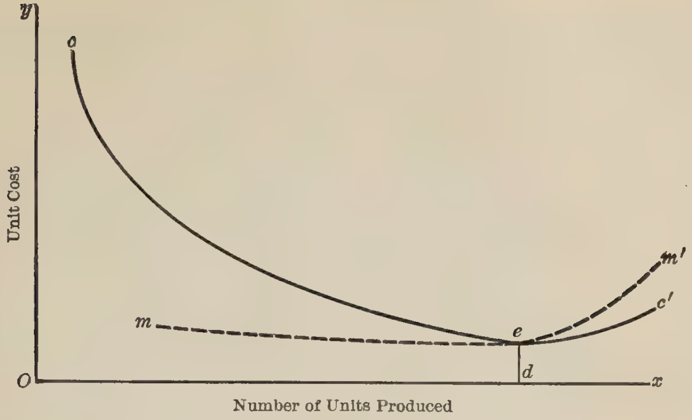


FIG. 36.

low point of the curve (see Figure 36) may not be reached till a very large plant is turning out a huge quantity of the product.

When the curves such as *cem'* (Figure 36) for the several producers are added together we get the curve *css'* shown in Figure 37.

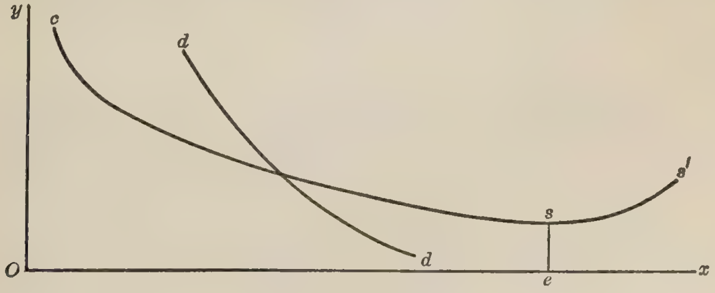


FIG. 37.

Remember that, by hypothesis, the industry is in the hands of a comparatively small number of large producers whose enterprises have not all reached the size of greatest efficiency or become fully engaged. The low point in the curve is thus very far to the right and represents a condition which has not yet been reached in fact. The demand curve is therefore to the left of this point and does not intersect the supply curve at all. The reader will recall that the curve *cs* is not a supply curve but a cost curve, showing, in its descending part, the minimum quantity which must be produced in order that all producers may have opportunity to break even at any given price. The intersection of this curve with the demand curve shows that at present the total product that can be sold is not enough to enable all the producers to operate at their lowest possible unit cost, but it tells us nothing as to the price at which the product will be offered for sale or the amount that will be exchanged. The purely potential supply curve off to the right is equally powerless. In short demand and supply under these conditions are not able to determine either the price or the quantity exchanged.

Of course the goods are being exchanged at a certain price or certain prices, but there is no settled market price. The situation corresponds to that condition, already noted, in the ordinary market, in which a few scattered sales at irregular prices may take place while the market is settling down to its equilibrium and before the market price has been established. Only now that condition, instead of being an accidental and quickly passing phase, is the prevailing state and may continue for months or years.

An unstable condition. Nevertheless we must emphasize what has been previously noted, that the state of decreasing costs is an unstable one. It means that the organization of the industry has not yet reached the point of maximum efficiency and lowest cost. There are more producers in the field than are necessary, and not all of their enterprises have attained to the most efficient size or become fully utilized. The point of maximum efficiency and lowest cost — the equilibrium point — can be attained only if the total product is further increased or if some of the producers drop out. Each producer realizes that if he could only increase his own output he could reduce his unit cost and so, as the result of both these causes, could greatly increase his profits. He is therefore tempted to make a drive for his competitor's trade by cutting the price. To any one of his competitors the loss of trade will be doubly disastrous, not only decreasing the amount of his sales but also increasing his unit cost. He must fight to hold his own, and the only way to meet his rival's attack is himself to lower the price.

Thus starts a price-cutting war which has no counterpart in agriculture, for example. Price cutting among farmers has a natural limit in the unwillingness of any farmer to sell for less than his marginal cost, which will be made higher, not lower, if he succeeds in getting trade from his competitors by increasing his own production. But when the law of decreasing costs prevails, there is no such limit. The prevailing price may be equal to the average cost, at which producers are just coming out even, but this will not deter any one of them from trying to increase his sales by offering to sell at a lower price.

And price cutting will not cease even when the price becomes actually lower than cost of production. The reader will do well to recall that it is the predominance of fixed charges in the total cost of production which gives rise to the condition of decreasing costs. And it is this also which induces competitive price cutting even after the price has fallen below any producer's average cost of production. Let us suppose that, at a given time, a manufacturer is turning out 1,000,000 units per month at an average cost of one dollar, made up of the following parts: fixed costs (interest on investment), 40 cents, overhead costs, 25 cents, particular running costs, 35 cents. Selling for less than one dollar is selling at a loss. However, if it is a question of selling below one dollar or not at all, the manufacturer will go on selling nevertheless. At a price of 65 cents for example he gets back his overhead and particular running costs, and has 5 cents toward fixed costs. But while he is thus losing 35 cents on each unit, this is better than losing the whole 40 cents of fixed costs which would be the result if he closed up. Fixed costs are "sunk costs"; they have been incurred once and for all; they cannot be escaped by going out of business. They are without power in determining the limit to falling prices when once competitive price cutting gets under way in an industry subject to decreasing costs.

And now we reach the surprising conclusion that if the prevailing price, as a result of competition, becomes actually lower than the cost of production, the price cutting is accelerated rather than retarded. Now each producer, selling at a loss, sees that the only way he can break even or make a profit is further to increase his sales until his cost of production comes down to the prevailing price. Before the motive was to increase profits; now it is to avoid losses. Desperation rules, and we have what is aptly called "cutthroat competition." The price is forced lower and lower, and as it goes down the quantity sold is increased, owing to the normal state of demand, but this increase cannot go far enough to bring equilibrium, since a price sufficient to induce buyers to take the necessary quantity would be lower than the lowest possible unit cost (see Figure 37).

It must not be inferred from the foregoing that the unstable condition of decreasing costs always and instantaneously leads to the régime of cutthroat competition. Here as elsewhere there is the discrepancy between the theoretical result under the ideal conditions assumed and the actual result in practice. There is today a growing tendency among business men to seek to avoid a war of cutthroat competition. If conditions are on the whole fairly satisfactory, the producers do not always attempt to squeeze out the last possible dollar of profit, even though their plants are not fully utilized and increased output might make possible a reduction of unit cost. Producers are loathe to start trouble and disorganize the industry by cutting prices. They are willing to let well enough alone. And when the majority so feel, they preach to the others the gospel of "live and let live," building up a sentiment in the industry against possible price cutting. Producers may even go a long ways toward making agreements to maintain prices. Since the condition of decreasing costs prevails only in those industries where production is naturally in the hands of a relatively few powerful producers, agreement or at least concerted action is not impossible. In such wise, an industry may go on for years in apparent harmony even though the conditions are such that many producers might lower costs by increasing their output.

But after all is said, the condition of decreasing costs remains in general an unstable one, likely sooner or later to degenerate into cutthroat competition leading to a new adjustment.

Equilibrium. The first step toward equilibrium is taken when one or more of the competitors gives up the fight and ceases production entirely. This enables the remaining producers to increase their respective products and so to lower their costs. As this development goes on, the dwindling company of survivors finally reaches the point where their enterprises have attained the size of maximum efficiency and are fully utilized. Then the cost curve for each takes the ascending shape and there is no longer motive to cut prices. This process is represented graphically by the gradual collapse of the curve *cSS′* (Figure 38) to the left as one by one the producers drop out, till finally it reaches the position of the dotted curve *css′*, in which the point *s*, at which maximum efficiency is first attained, is to the left of the demand curve, and there is intersection of the curve of demand with either the constant cost curve or the supply curve. Equilibrium has been reached, and a normal condition of constant or increasing costs henceforth prevails.

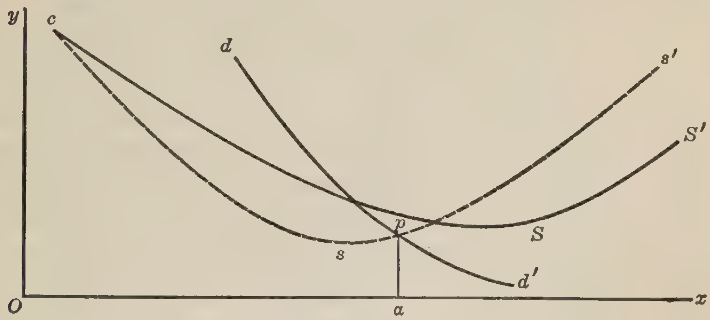


FIG. 38.

It will be noted that equilibrium is attained only by reducing the number of the producers, the reason being that, until concerns have reached the size of maximum efficiency and become fully utilized, production is cheaper the smaller the number of the producers. How far this reduction of producing units will have to go before equilibrium is reached depends upon the conditions of the particular industry. In the extreme case the result may not come until all producers but one have been eliminated. Decreasing cost may thus lead through cutthroat competition to monopoly. This is what happens generally in the street railway business, in the local telephone business, in the business of providing water, gas, and electricity to a city, etc. One producer can furnish the entire service more cheaply than two or more. In the majority of industries equilibrium comes before absolute monopoly is reached, and we have a few great corporations or trusts, each one about as large as can be successfully operated and with its plant fully utilized, though a few small competitors will often be found content to take the business left by the great concerns. Several industries in America are today in this state, as, for example, petroleum refining, sugar refining, cigarette manufacture, automobile manufacture, etc. Frequently the parties to cutthroat competition do not wait for the final result but come together and either by voluntary combinations reduce the number of the competing producers or by mutual agreements put a limit to price cutting.

Qualifications. That the conclusions at which we have now arrived are broad statements of tendency, of the normal results which follow from certain causes under specified conditions, and that in actual business more or less wide departure from these results will be found when the conditions are not precisely fulfilled will be clear to the reader who has grasped the nature of all scientific generalization. He will realize that in the complex of innumerable forces which act and counteract in the world of practical affairs, it is impossible to include all in any general statement, and he will be prepared to meet cases in which the adjustments between cost and price here formulated are not perfectly accomplished. Two circumstances in particular make it difficult or impossible ever to take account of all the factors in any given price problem. One is the constant state of change which, as we have already had occasion to observe, is normal to industry. Before any relation has time to work itself out, the basic conditions are likely to change, thus causing new readjustment of causes and effects. In the second place the whole adjustment depends upon human judgment. The entrepreneur proposes neither to sell goods nor to produce more unless he can get back at least what they will cost, and he dare not stand out for much more lest he lose his trade to his competitors. But no one can know in advance exactly what it will cost to produce any commodity. The entrepreneurs have to make their decisions, taking account of all the present circumstances and estimating the future, according to their best judgment. Their judgment, though on the whole surprisingly keen, will never be perfect and it may sometimes lead them far astray. The personal element is thus always a more or less uncertain factor in the problem.

Nevertheless, in spite of all uncertainties, we do have a generalization which describes the normal operation of those forces which determine cost of production and supply and their relation to price. On the whole and in the long run, supply does conform to cost of production in the manner thus laid down. The uncertain factors are not sufficient to destroy the fundamental relation and to make the result the mere whim of chance. Occasional wide discrepancies can generally be recognized as exceptional, and very frequently it is possible to discover the peculiar circumstances to which they owe their existence. And we are generally safe in branding such discrepancies as " abnormal " and looking for restoration of the normal.

Cost of production and price changes. The reader will by this time probably have passed beyond the danger of confusing supply and the curves illustrating production costs and supply with historical records of changes in cost of production and price. For example he will not be guilty of the error of thinking that the normal ascending supply curve means that increased quantities of the good can be obtained from time to time only by raising the price. Of course a single supply curve tells us nothing about this. Supply relates to a particular time and, in conjunction with demand, determines the price and the quantity exchanged at that time. A day, a week, a month, a year later there will be a different supply (as also a different demand) and a different price and a different quantity exchanged. The quantity may be larger and the price lower, though the supply curve is still an ascending curve. There is no mystery about this.[1]

Changes in the price of any good are caused by changes in demand or supply or both, or sometimes directly by changes in the cost of production. This was demonstrated in Chapter XIV, where however nothing was said as to the ulterior causes which might be responsible for the changes in demand or supply. Having now learned how supply depends upon cost of production, we are able not only to explain the position and shape of the supply curve at any given time (which has been the main goal of the present chapter), but also to see more clearly the causes which in the course of time may increase or decrease supply; *i.e.*, shift the supply curve to the right or left, with consequent results upon price.

In agriculture for example the law of increasing costs prescribes an ascending supply curve and it also indicates that, other things being equal, the supply curve will not materially change its position in the course of time. Increase in demand, as population grows, thus causes prices to rise. However other things do not always remain equal. There have been, in the history of agriculture, important discoveries of new machines and new processes. The three-field system of cultivation employed on the medieval manor was not the most efficient. Later discoveries in the scientific rotation of crops and the scientific use of fertilizers have made possible an increased product at a lower cost. Modern labor-saving agricultural machinery has had the same effect. Thus invention from time to time pushes the supply curve to the right, tending to counteract the constant increase of demand. We thus have the mechanism through which growing population and diminishing returns in agriculture have for centuries been more or less steadily increasing the prices of agricultural products, checked from time to time by invention of more efficient processes or mechanical appliances.

On the other hand there are products, cotton and woollen cloth for example, which for long periods exhibit falling prices in company with enormous increase in the quantities produced and sold. This is often due to the same cause which from time to time increases the supply of agricultural products; namely, invention. In the textile industries the epoch-making machinery, whose history we studied in an earlier chapter, so enormously lowered the unit cost of production that this change more than counteracted all increase in demand due to the growth of population. The result has been lower prices and extraordinary increase in output, as illustrated in Figure 39.

---

[1] See Chapter XIV, where the whole matter was investigated in detail.

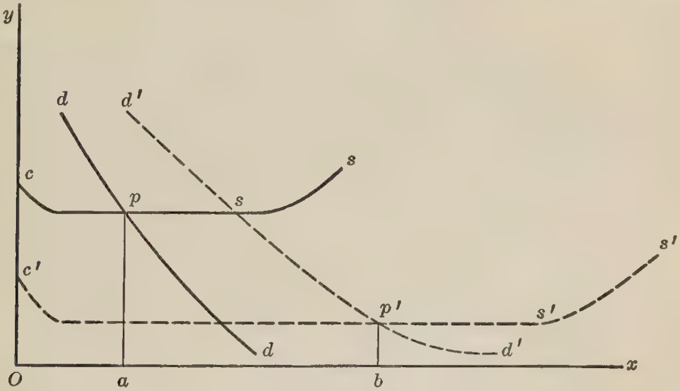


FIG. 39.

Finally in those industries subject to decreasing costs we have seen how the existing unstable situation can be brought to equilibrium only by a shift of the constant cost curve or the supply curve far enough to the left to make possible an intersection with the curve of demand and enable the producers to operate at maximum efficiency.

Competition. In an earlier chapter (Chapter V) attention was called to competition as one of the foundation stones of the modern economic organization. We are now in a position better to understand how competition works in the determination of price and as the regulator of industry. When competition has opportunity to work, it sees to it that no producer can long remain in business unless he is able to turn out goods at prices which neither exceed the marginal utility of the goods to the consumer nor fall below his own cost of production. The producer who has missed the public taste is ultimately eliminated, as is also the one who is not efficient enough to keep his costs as low as his competitors. The consumer, on the other hand, cannot have things to enjoy unless he is willing to pay for them prices normally equal to their cost of production, and he will not normally have to pay much more than that.

Competition is not popular with the ordinary business man. No one likes the authority that regulates him. Business men are constantly seeking to eliminate or reduce competition as it affects them. The protective tariff is advocated in order to remove the competition of the foreign producers. The business man is no more friendly to the competition of his fellow citizens in other states or towns of his own country. The device of the protective tariff is not available, but he seeks the same end by trying to persuade his community to favor him as against the outsider. "Buy at home," "Patronize home industry," "Put your money in circulation where it will help your own community" are examples of the slogans used in this sort of campaign. The trolley cars carry signs reading "Patronize your neighborhood grocer." Back of all such persuasion is the desire to counteract the natural tendency of people to buy in the best market, which is the foundation of competition. It represents the struggle of producers to escape the controlling hand of competition. It is for the same purpose that producers come together and make agreements to eliminate competition between themselves. Thus there arise combinations of various sorts. And finally there may come monopoly, in which case competition between producers has entirely ceased.

Combination and monopoly. Competition assures to the consumers prices tending in the long run to be equal to cost. No such assurance is present under combination and monopoly. The causes affecting supply and price under such conditions will be studied in the next chapter, and in a still later chapter we shall have to inquire into the special problems which arise when consumers are thus deprived of the protecting influence of competition.

CHAPTER XVIII

MONOPOLY PRICES

Classification of prices. Thus far in Part II attention has been directed primarily to competitive prices. It remains to investigate in the present chapter the forces which determine non-competitive prices, meaning all prices where there is not competition of sellers. Non-competitive prices may, for purposes of study, be divided into two groups, which finally gives us the following classification of all prices:

1. Competitive prices (competition of sellers).
2. Non-competitive prices (no competition of sellers).
    - *a.* Monopoly prices (competition of buyers but not of sellers).
    - *b.* Higgling prices (competition of neither buyers nor sellers).

Higgling prices. Where there is only one seller (or sellers acting together in a combination) and only one buyer (or combination of buyers), we can go only a certain way in developing general principles governing the price and the quantity sold. Seller and buyer must come together and by a process of higgling or bargaining determine first whether a sale is possible. This requires that the lowest price which the seller will accept be not higher than the highest price the buyer is willing to pay and also that an agreement upon the quantity is possible. Having determined that there is possibility of terms mutually satisfactory, the process of bargaining must finally fix the price and the quantity to be exchanged. Such transactions, though requiring our notice in order to complete the analysis of price, are of comparatively rare occurrence and of slight economic importance. They will require no further study in this connection.

Monopoly. The subject of monopoly price is of far greater significance, since there are numerous commodities and services of great public importance whose production and sale are in the hands of sole producers or combinations. There are monopolies controlling the sources of certain products of nature, such as the Carlsbad mineral waters, the diamonds of Africa, or the anthracite coal of Pennsylvania. The possessor of any unique non-reproducible article, such as a famous work of art, is of course in a monopoly position. A famous actor or musician has a monopoly in the personal services which the public demands of him alone. The United States government has a monopoly of the service of carrying letters. Patents and copyrights give monopolies to inventors and authors, and city governments grant monopolies to corporations in the form of exclusive franchises to operate street railways, to furnish water, gas, and electricity, etc. Finally there are those great corporations and combinations, in fields in which the advantages of large scale production and combination are especially strong, which have come to positions at least approaching monopoly in their particular fields.

Anticipating the more thorough analysis of the nature of monopoly in its various forms which will occupy our attention in a later chapter, the examples cited here are sufficient to indicate the very broad field in which monopoly may exist and the importance of the subject of monopoly price.

How the monopolist determines the price. All that has been learned of demand is applicable to the determination of monopoly price exactly as of competitive price. The monopolist has no control over demand. His position of privilege relates to the seller's side of the market only. He may take advantage of his position in one of two ways, either by fixing the price at which he will sell or by fixing the quantity he will sell. In case he takes the first course, demand will determine the quantity he can sell. In the other case, demand will determine the price. These principles and the considerations which guide the monopolist in making his decisions as to what price to charge or what quantity to sell may be made clear by use of the graphical aids with which the reader is now familiar.

Let us suppose that a manufacturer owns the patent for a certain article, giving him a monopoly in its manufacture and sale. The relation between his unit cost of production and the quantity produced will be such as is shown by the curve *css′* in Figure 40. This is the ordinary cost curve with which we are now familiar, the descending part, *cs*, showing average cost, the ascending part, *ss′*, showing marginal cost.

Increasing costs. Now let the demand for this product be indicated by an ordinary demand curve, *dd′*. If competition were in force, the intersection of these curves at *p* would determine the price *pk* and the quantity exchanged *ok*. But the monopolist does not have to accept this price; he will seek to fix whatever price will give him the greatest possible gain. Obviously he will not set a price below *pk*; that would mean a loss on the last units produced. His price, if not *pk*, will be something higher, let us say *oe*. Having set the price *oe*, the quantity he can sell is determined by the demand curve, intersecting a horizontal line from *e*. The quantity is *of*. By charging a price higher than his marginal cost, the monopolist receives more for each unit of his product but sells fewer units.

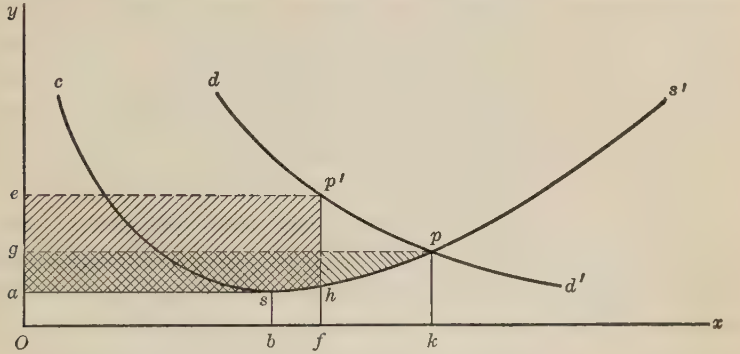


FIG. 40.

The relative effects of these two prices upon his gain are disclosed by the diagram. At the price *pk*, the total receipts from sale of the product are represented by the area *ogpk*, since this area is the product of *pk*, the price, by *ok*, the quantity sold. The total cost of producing this quantity is represented by the area *oaspk*, being the sum of the area *oasb*, which is the cost of *ob* units at an average cost of *sb*, and the area *bspk*, which is the cost of producing the quantity *bk* subject to rising marginal cost. The difference between the total receipts and the total costs is the gain and is represented by the shaded area *agps*. By putting the price at *p′f*, the gain is the other shaded area *aep′h*. It is this area which the monopolist will seek to make a maximum through his control of the price.

The foregoing analysis covers the ordinary cases of monopoly price when production is subject to increasing costs. The monopolist will almost certainly set a price higher than competition would have determined; that is, higher than his marginal cost. But, while he can increase the gain per unit by raising the price, he will at the same time cut down the number of units sold. The one effect tends to increase his total gain, the other to reduce it; somewhere therefore there must be the price most favorable to him because it makes his gain the maximum. Whatever price he fixes will represent the closest approximation he is able to make to this price.

Whether the monopoly price is fixed high or low depends mainly upon the elasticity of the purchase. If purchase is elastic, indicated graphically by a gently sloping demand curve, there is relatively great loss of sales from a high price, which may soon overbalance the larger gain per unit. In such cases the monopoly price may be comparatively low. On the other hand, when purchase is inelastic (shown by a steep curve), the price may be put very high without diminishing sales enough to neutralize the large gain per unit.

Decreasing costs. Let us consider now the case of decreasing costs, which would occur if the total quantity which buyers would take at any profitable price is less than enough to enable the monopolist's plant to operate at maximum efficiency. An example would be a railroad which did not have enough business to keep its tracks and terminals and rolling stock busy. Graphically this situation is illustrated by Figure 41. Only by producing the quantity *oa* can the monopolist have his plant fully occupied and bring his average cost down to the minimum. But buyers will not take

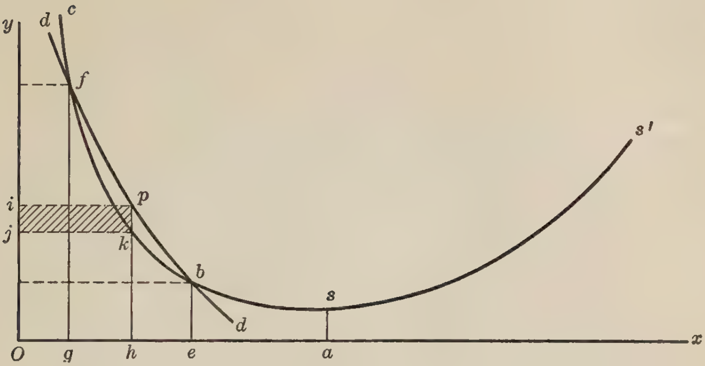


FIG. 41.

that quantity except at a price less than the minimum cost, perhaps not at all. The monopolist must therefore be content to have his plant only partially utilized; that is, to operate under conditions of decreasing costs.

Now it is evident at once that the two points at which the demand curve cuts the curve of average cost determine the limits within which the price will be set. A price either less than *be* or greater than *gf* will be lower than the average cost of the goods which could be sold. Somewhere between these limits the monopolist will fix his price, guided as always by the desire to obtain the maximum gain. For example if the price is fixed at *oi*, the demand will determine the quantity that can be sold, *oh*. The total gain will be the area *jipk*, and it is this area which the monopolist will seek to make a maximum.

As to the rest, this case is similar to that of increasing costs. According to his best judgment and experience, the monopolist fixes a price at which the product of the gain per unit by the number of units sold will be as great as possible. And here also the height to which the price may go depends primarily upon the elasticity of purchase.

We do not need to consider the third cost condition, constant costs, since it is hardly conceivable that a single producer could operate for any considerable time under conditions of constant cost, implying that he had neither fixed charges nor overhead expenses.

When the monopolist determines the quantity to be sold. While theoretically the monopolist may, instead of fixing his price, determine the quantity he will sell and let the demand then settle the price, this choice is scarcely practicable except when it is possible to sell the whole product to a single buyer through the medium of an auction. In that case, the monopolist puts up a certain quantity, having decided that that amount can be sold at the price most favorable to him. Through the bidding of the rival buyers, the price is then determined by demand.

Non-reproducible goods. Thus far in this chapter we have had in mind the price of reproducible goods. Of course there may be a monopoly of goods not capable of reproduction, as when a person is in possession of a famous painting or all the copies of a coveted first edition or all existing samples of a rare postage stamp. Here the monopolist will, in selling, seek merely the maximum total receipts, there being no question of cost. Either he will set the highest price at which he thinks he can dispose of the number of units he wants to sell or he will put up a definite quantity for sale at auction to the highest bidder.

High and low monopoly prices. There is a common impression that monopoly price is always a very high price, but this is not true. Some monopolists do charge very high prices, but there are others which find it more profitable to make their prices comparatively low and to obtain the maximum gain through large sales. A monopoly which controls an article of very elastic purchase — a luxury or a costly convenience — must be careful not to put its price too high, lest its sales fall off and its gains decline or vanish. Wise selfishness may in such cases dictate a fairly reasonable price, not much above the cost of production. On the other hand the fortunate monopolist who controls a necessity of life, like salt, or some other article of inelastic purchase, like tobacco, may charge extremely high prices without seriously cutting into his sales and so may reap handsome monopoly gains.

Limitations upon monopoly price: substitutes. We have obtained a view of the forces which guide the monopolist in determining his selling policy and fixing his selling price. It must not be presumed however that every monopolist is free thus to determine the price of his product in accordance with his own unrestricted wishes. There are indeed some very potent restrictions upon the price-fixing power of the monopolist. First of all we note the pressure of substitutes. Few commodities or services are absolutely indispensable. A patented article is very often simply a superior device for doing what has heretofore been accomplished in some other way. Too grasping a price policy may lead people to return to the former device. The waters of a monopolized mineral spring may be much desired; but if the price is considered unreasonable, many consumers will find that they can get on with another brand. Diamonds are in competition with other jewels, and anthracite coal faces the competition of bituminous coal and other fuels. The local gas company may, by exorbitant rates, drive people to use oil lamps and to substitute coal stoves for their gas ranges. Even the popular singer or actor has to remember that there are others and that there is a limit to what the people will pay for his services. Thus the ever present possibility that "something just as good" may appear puts a very real check upon the greed of many a monopolist.

Potential competition. It must in the second place be recognized that the position of few monopolists is absolutely impregnable. The complete absence of competition which marks perfect monopoly is often lacking, even though the one leading producer so far overtops all others as to make his position predominant and give him virtually all the privileges of monopoly. Frequently his position is maintained only because his price policy is reasonable. Let him begin charging excessive prices, and some of his small and hitherto negligible competitors may be induced to enlarge their operations and challenge his supremacy. And even when there are no competitors in the field, this is often not because competition is impossible but because the inducements are not sufficient to lead a new producer to embark on the long uphill fight against the well-entrenched monopoly. However if the price of the product is high enough and the monopoly gains great enough, someone is likely sooner or later to make a bid for a share, and so competition starts. Competition which, though actually not present or so weak as to be negligible, is always likely to be aroused is called *potential competition.* Nearly every one of the great " trusts " or combinations in the United States which has followed too grasping a price policy has sooner or later found itself facing some unexpected competitor, and the danger of arousing this potential competition imposes caution and restraint upon many monopolies.

Risk of legal interference. Combinations and monopolies are never popular with the consuming public, and one that charges excessive prices is very likely to attract public attention. There arises a popular clamor against the monopoly and a demand for government interference to limit the price or in other ways to regulate the monopoly's business. Politicians are always willing to put themselves at the head of such a movement and to lead the drive against " the trusts " or " the interests." There have been any number of cases in America of interference with industrial monopolies by Congress or the state legislatures. The railroads have in this way been very seriously restricted. They are no longer free to fix their rates solely according to their own desires, and the conduct of their business operations is supervised and regulated in a multitude of details. This was their punishment, doubtless only partly deserved, for the excessive rates and the callous disregard of the consuming public which marked the policy of many railroad managers in the palmy days of railway monopoly. Similarly the Standard Oil Company, the great meat packing corporations, and numerous other concerns that have or are supposed to have monopoly powers, have been subjected to Congressional and legislative regulation or at least investigation. The trolley companies, gas and electric companies, water companies, etc., have in the United States generally been deprived of the unlimited power to determine their rates and have had to submit to government control not only of their rates but also of the character of the services rendered by them.

Legal interference has thus already made great strides, and the last thing desired by any monopoly which has thus far escaped is unfavorable public attention and legal investigation and control. The wise monopoly manager is thus willing to restrain his inclination to raise prices unduly.

Conclusion. We conclude therefore that monopolies in general do not have the unlimited power over prices which many persons suppose and that monopoly prices are not always so excessive as common opinion holds. Nevertheless there can be no doubt that monopoly is not generally in the interest of the public. As we have learned, competition in the long run ensures that the people get the kinds of commodities and services they want at prices not far from the cost of production. Monopoly makes no such promise. The monopoly price will almost always be somewhat higher, and it may be very much higher, than cost of production, and the monopoly is not under so strong inducement to produce the kind of goods the people want.

The seriousness of the situation is much dependent upon whether the monopoly controlled commodity is a luxury or a necessity. Monopoly control of a luxury is no great menace. As we have learned, the monopoly will, in its own interest, generally refrain from charging an exorbitant price. And if its price is excessive, the people can do without this particular article. Control of a necessity of life by monopoly is a far more serious matter. The monopoly is here able to charge a high price, and the people have little opportunity to escape its payment. In general the public interest is better served when production is carried on under the régime of competition.

There are certain exceptions, some self-evident, others not so obvious. Personal monopolies, such as those enjoyed by the popular singer or actor, are inevitable and can do the public little harm. The monopoly granted by the patent or copyright is generally in the public interest. And there are certain lines of industry, such as railroading, telephone, telegraph, the furnishing of water, gas, electricity, etc., which from their economic nature appear destined to be monopolies whether or no. In such cases it is futile to put our trust in competition which cannot endure. Monopoly must be accepted and made the best of. This topic will present itself for our further study in a later chapter.

Public opinion today generally holds that monopoly, particularly in the necessities of life, should so far as possible be avoided in the public interest and that when monopoly is inevitable, government regulation should be relied upon to ensure to the public good service and reasonable prices. For this general opinion there is firm foundation in economic principles.

EXERCISES

1. Draw diagrams showing the comparative effects of elastic and inelastic purchase upon monopoly price, (*a*) under conditions of increasing costs, (*b*) under conditions of decreasing costs.

2. A government has for sale 10,000 army tents of a unique pattern, which it puts up for sale at auction. Draw a diagram illustrating the way in which the price would be determined.

3. The price of a ticket to the Yale-Harvard football game is $5.00. The Bowl seats approximately 79,000, not including those admitted free. The number of persons who try to secure tickets each year is much in excess of the seating capacity of the Bowl.
   - (*a*) Draw the demand curve illustrating this situation.
   - (*b*) Explain how the price and the quantity of seats sold are determined.
   - (*c*) Suppose the Athletic Association should decide to obtain the greatest possible receipts from this game. State the result, and illustrate by your diagram.
   - (*d*) According to your diagram, how many tickets could presumably be sold at $5.00 if the capacity of the Bowl were unlimited?

4. The management of a theatre with a seating capacity of 3,000 fixes a uniform price of $2.00 per seat for a certain performance.
   - (*a*) On the night of the performance, only half of the seats are taken. Draw a diagram illustrating this situation.
   - (*b*) According to your diagram, what price would have just "sold out the house"?

5. On the next night, the price still being $2.00, every seat in the theatre of the previous exercise is taken, and there is a crowd outside trying to buy tickets.
   - (*a*) Draw a diagram illustrating this situation.
   - (*b*) According to your diagram, how many tickets could have been sold, if the theatre had been large enough?
   - (*c*) According to your diagram, what price could the management have charged and still have sold every seat?

CHAPTER XIX

THE PRESENT VALUE OF FUTURE INCOME

Income the goal of economic activity. Man endures irksome labor in order to obtain income for himself and his dependents. In a society of freemen it is on this universal desire for income that reliance is placed to get the work done which is required to supply the members of the society with the necessaries, the comforts, and the luxuries of human existence. The cultivation of the land, the cutting of the forests, the exploitation of the mines, the working up of raw materials into forms fit for the satisfaction of human wants, the construction of systems of transportation and of agencies for marketing products are all the work of men who are inspired in their labor by the expectation of sharing in the streams of benefits which flow from these activities. Thus the desire for income is the mainspring of all economic activity. It is in terms of income that the effect of economic forces upon human welfare is measured.

Capital the means to the end.[1] We have learned that all production requires capital. Hence to obtain income one must gain control of the capital goods with which it is produced, either through outright ownership and possession of these goods or through the acquisition of partial property rights in them. In a primitive society this effort to secure capital takes the direct form of devising and constructing the weapons, tools, buildings, and other instruments by means of which the productivity of man's labor is multiplied and his enjoyable income increased. Under such simple conditions therefore it is customary for the same man both to bear the final costs of capital — the labor of making the instruments and of keeping them in repair — and to consume its benefits. Neither capital nor income is frequently bought and sold; hence neither acquires an exchange value and there is no tendency to dissociate capital from income and regard it as an end in itself. Capital is viewed in its proper light as a means toward an end, and, like other means, each unit of capital is more or less prized according to its efficiency in achieving the end in view; that is, its power to produce income for its owner.

---

[1] In the present discussion we shall use the term capital in the broadest of the several meanings which have become irretrievably attached to it; *i.e.*, as a synonym for wealth, thus including land as well as man-made instruments. The term is generally so used in discussions of investment and when it is desired to give prominence to the distinction between wealth — the fund — and income — the flow. Cf. Chapter I and the footnote on page 106 (Chapter VI).

In the social order of today there are to be found many instances of this direct and visible relation between capital and income. Men sometimes gather materials and with their own labor construct some simple capital instrument whose benefits will accrue to themselves directly. In all such cases the relation of capital to income is clearly grasped; one sees that it is the expected income which bestows on capital its importance, and that the owner's appraisal of the capital is but a reflection of his evaluation of the benefit which he hopes to obtain from it. An article of capital from which the owner fails to obtain benefits will be considered worthless regardless of the amount of labor expended upon its construction. On the other hand one which is very productive of income will be highly prized even when it has been obtained at small labor cost.

Property and capital. But these simple cases are not typical of present-day society. So complex has our economic structure become that we often lose sight of the fact that capital has no significance when divorced from the benefits which it produces. Chief among the attributes of our social system which have this effect is the fact that the ownership of capital has been subdivided into a large variety of property rights which are bought and sold in great numbers. These property rights are usually represented by certificates of different kinds — stocks, bonds, mortgages, contracts, and the like — which frequently have stamped upon them some nominal valuation (for example, the $100 " par value " on a stock or bond) which misleads the owner into the belief that these certificates have a " normal " or " true value " in themselves. The development of many agencies with the purpose of facilitating trading in these property rights, such as banks, brokerage and investment houses, stock exchanges, and so on has aided in beclouding the true relation of capital to income.

When certificates of property are bought and sold in great numbers, they fall into the hands of people who have neither knowledge of the capital of which they are evidence of ownership nor interest in the direct income which that capital produces. The owner of stock in a railroad for example need not, and generally does not, know anything at all about the technique of the capital instruments which make up a railroad system, nor need he consciously share in the enjoyment of the benefit (transportation) whose production is the reason for the existence of those particular capital instruments. He may merely look at the par value of his stock and conclude that for some reason this bit of paper has a true value of that amount. Even if he is inclined to think a little farther than this, he will probably be content with the explanation that it is the particular bit of railroad equipment to which his stock gives him title of ownership which has a "real" value of $100. He would be at a loss to explain why these concrete capital instruments are worth $100, unless indeed he had recourse to the customary explanation that they probably cost that much to produce. In general it is true that the average man has very vague and incorrect notions of the way in which the value of capital and of property rights in capital is determined. It is the purpose of this chapter to show that the value of capital and property rights depends fundamentally upon the value of the income which the capital produces and to explain the process by which this factor enters into the determination of the value of capital.

Property valuation. Let us first consider the relation of the value of property to the value of capital. As we have learned, property is but a right to the future benefits of wealth (capital). Property does not produce these benefits but merely determines who shall get them. The benefits are produced by the capital which underlies the property; it is the function of the property right to divide this income among the owners and part owners of the capital and to debar others from sharing in it. It is apparent therefore that the principal reason for the existence of the certificate of property is that it affords a convenient method of transferring ownership in capital from one man to another and of subdividing the ownership so that many may share in the income produced by a given set of capital instruments. These certificates are bought and sold as representatives of the valuable thing (capital) in which they transfer ownership; hence the action of the buyers and sellers in the market will result in giving them the same value that in the opinion of these buyers and sellers should be placed on the capital which they represent. If for example a manufacturing concern has issued one thousand shares of stock and these shares sell in the market for $150 each, there can be but one reason for this; namely, that in the opinion of investors all the assets of the concern are worth $150,000 over and above the amounts owed to the creditors. The fact that the stock has stamped upon it a nominal value of $100 or $200 or $50 or any other figure, will not affect its market value as long as the capital which the stock represents is believed to be worth $150,000. The reader will recall that the subject of valuation of capital stock was investigated in more detail in Chapter VI.

Present worth. The real problem therefore is to discover how the value of capital goods is determined. Here we must repeat that capital is desired only because of the income it produces and is bought only as a means of obtaining this income. Let us suppose that a given machine will last ten years and produce a net income of $100 a year. If this machine is offered for sale to the owner of a factory, he will consider that by owning it he can increase his income during the next ten years by the amount of $100 a year, or $1,000 in all. Whatever price he is willing to pay for the machine will be a measure of the present value he puts upon this thousand dollars, distributed as it will be through the ten years to come. Or, to change the illustration, we may suppose that an acre of land is offered for sale to a man who has reason to believe that it can be made to produce a net income worth $10 a year forever. In this case too the price offered by the buyer for the land will be his estimate of the *present worth* of an annual income of $10 which continues indefinitely. Or again an investor may be offered a bond which will be redeemed for $100 at the end of five years and in the meantime pay $5 a year interest. Such a bond would be viewed as an opportunity to receive a series of money payments during the next five years: $5 one year from now, another $5 two years from now, another $5 three years from now, $5 more four years from now, and a final payment of $105 after five years have elapsed. The amount of money that the investor is willing to part with today in order to buy the bond must be an indication of what he considers this series of future payments to be worth at the present moment. This is the essential thing in calculating the marginal utility of any capital good or of any property right in capital goods. It amounts to placing a present value on an expected future income of a certain size and form and duration.

Present goods more desirable than future goods. But by what process can we link present and future values together? It is a well-known fact that men do not consider a sum of money which will accrue in the future to be worth its face value now. The investor referred to in the preceding illustration would not calculate the present value of the bond by simply adding together the different payments which will accrue in the future to its owner. Such a calculation would give the bond a present value of $125; and the buyer paying this price for it would eventually get back just the amount of money that he paid, though he would have to wait five years for most of it. To change the illustration, if one offered for sale a promissory note which would pay its owner $100 twelve months from today, he would find no one willing to pay $100 for it now. Nor would this reluctance to buy this future income at its full face value in present money be due only to risk of loss. If the $100 were locked up so safely that all doubt regarding its security were removed, still no one would be willing to pay out $100 now and be content to get it back after twelve months had passed. Human nature is such that men will not exchange present dollars for future dollars on terms of equality, or a given amount of present goods for the same amount of future goods, or present income for future income of the same amount. Today's dollar and the dollar of next year do not have the same present value in the mind of the ordinary man. The present value of the future dollar is less than that of today's dollar; hence sums of money which accrue in the future will be bought only if they can be had for less than their face value. And the price that will be paid for capital which gives its owner the right to receive income in the future will be lower than the value of this future income when it accrues.

Interest. If asked why he will not pay $100 now for the right to receive $100 one year hence, the practical man will probably answer that he is unwilling to lose interest on his money. Because of the existence of interest, it lies in the power of any man who has present wealth to exchange for future wealth to make the exchange on terms which will bring him a future return of larger value than his present payment, and as long as this opportunity to gain interest on the money they lend is open to all men, no one will consent to pay its full future value in present money when buying income. *Interest* may be defined as *the premium which present goods command over similar goods due in the future.* The man who lends $100 for one year at five per cent interest exchanges $100 worth of present goods for $105 worth of next year's goods, the premium of $\frac{5}{100}$ being added to the future goods to give them a present value equal to that of the present goods which the lender surrenders. As the interest rate rises and falls the premium which must be added to make these two amounts — one present, the other future — of equal present value rises and falls accordingly. It is, therefore, the *rate of interest* which determines how much larger the future return must be than the present sacrifice in order to cause the exchange of the two to occur.

We need not at this point attempt to explain why interest exists or how the rate of interest is determined; these are problems to be solved later in our study. For our present purpose it suffices to accept the existence of interest as a feature of the present business world. Our immediate concern is to show that the rate of interest supplies the link between future and present value which enables us to translate the one into the other. Given the rate of interest, it is a simple matter to calculate the equivalent future value of any amount of present money. When the interest rate is five per cent, for example, \$100 invested now will be worth \$105; *i.e.*, \$100 × 1.05, one year from now; \$110.25; *i.e.*, $\$100 \times (1.05)^2$, two years from now; \$115.76; *i.e.*, $\$100 \times (1.05)^3$, three years from now; and so on. This translation of present into future values through the medium of the interest rate may be reduced to a formula, as follows:

$$A_n = P(1 + r)^n$$

in which:

A represents the future value, any number of years hence;
P represents the present value, or "principal";
r represents the rate of interest;
n represents the number of years.[1]

This same formula may be adapted to the reverse problem of determining the present worth of a given future value by merely rearranging its terms, as follows:

$$P = A_n \div (1 + r)^n$$

the symbols having the same meaning as in the preceding formula. To illustrate, assuming again an interest rate of five per cent, \$100 due one year hence is worth now \$95.23; *i.e.*, \$100 ÷ 1.05; if due two years hence, it is worth now \$90.70; *i.e.*, $\$100 \div (1.05)^2$; if due three years hence, it is worth now \$86.31; *i.e.*, $\$100 \div (1.05)^3$; and so on.[2]

---

[1] The demonstration of this formula is as follows:

At the end of the first year, the amount (A_1) is the sum of the present value, or principal, and one year's interest; that is,

$$A_1 = P + Pr = P(1 + r).$$

That is, the amount at the end of the first year is the principal multiplied by $(1 + r)$.

This amount, $P(1 + r)$, is the new principal for the second year. The amount at the end of the second year (A_2) will be found in exactly the same way as for the first year; that is, by multiplying the principal at the beginning of the year by the factor $(1 + r)$. Thus

$$\begin{aligned} A_2 &= P(1 + r)(1 + r) \\ &= P(1 + r)^2. \end{aligned}$$

In the same way, we find at the end of the third year the amount is $P(1 + r)^3$, at the end of the fourth year, $P(1 + r)^4$, and so on for any number of years.

[2] The examples in the text assume that the interest rate is so much *per annum* and that interest is "compounded" annually. The formula holds however for any unit of time. All that is required is that the rate of interest be stated with respect to the same period as is used in compounding. For example, it is required to find the amount of a loan of \$5,000 for two years at the rate of one per cent a month, interest compounded semi-annually. Since the interest is compounded every half

The discount process. Although it is possible by means of the interest rate either to calculate future values from present or present values from future, and although in the business world there is continual use for both of these calculations, it is the latter process which has the greater importance. This is true for the reason stated at the beginning of this chapter; namely, that the economic activity of mankind is directed toward the obtaining of income and all income lies in the future. The valuation of capital instruments and property rights is incidental to this struggle for income, being a product of the incessant buying and selling of the means by which income is produced. Hence the calculation of the present values of future incomes of every conceivable variety as regards size and duration is one of the commonest events in the business world. The name given to this calculation of present values of future incomes is the *discount process.*

Summarizing the preceding discussion, we may say that this process involves these two fundamentals: (1) a knowledge of the probable future value of the income, (2) a knowledge of the interest rate. The first of these fundamentals is the basis of the calculation — the future value whose present worth it is desired to obtain. The second supplies the mechanism for determining this present worth; the process being to divide the future value by one plus the rate of interest raised to a power equal to the number of years which must elapse before the future incomes accrues.

Capital and income. It will be observed that the relation of capital to income may be looked at from either of two points of view; the point of view of the physical relationship and that of the value relationship. From the physical point of view it is capital which produces income, an orchard produces a certain number of barrels of apples, a factory, so many pairs of shoes or other units of output, a house, so many years' shelter, and so on.

But the value relationship is the opposite of this. The value of the capital does not produce the value of the income. The

---

year, we must take a half year as the unit of time. Then we must call the rate of interest, instead of one per cent per month, six per cent per half year. Instead of saying the loan is for two years, we must say it is for four half years. Then our formula becomes $A = 5{,}000\ (1.06)^4$ and, when we solve, we find that $A = \$6{,}312.38 +$.

value of the apples gathered from an orchard will be established in the market in which the apples are sold after they have been produced. The price of shoes will also be determined in the market by the forces of demand and supply which operate when the shoes are sold. So it is with the product of any capital instrument; its value will be governed by the laws of market price as explained in the preceding chapters. Setting a value upon the capital will not avail to dictate the terms upon which its product can be sold. A given set of capital instruments may be very productive, in the sense that it turns out a large stream of products, but if these products cannot be sold at profitable prices, the value of the capital instruments will be very low, or even non-existent. This is true of capital when the demand for its products has declined or disappeared. The abandoned farms, factories, and railroads which one occasionally sees bear testimony to the fact that the value of capital goods is derived from that of their income, just as the worthless stocks and bonds which accumulate in the hands of ignorant investors are painful evidence that property rights have no value in themselves.

This two-sided relationship between capital and income is pictured by the following diagram:

$$\begin{tikzcd}
\text{Capital goods} \arrow[r] & \text{Flow of benefits (income)} \arrow[d] \\
\text{Value of capital} & \text{Value of income} \arrow[l]
\end{tikzcd}$$

This shows that capital in the physical sense produces income; that income derives its value from the conditions governing the market at the time it is sold or enjoyed; that the estimated amount of future value of income governs the value of the physical capital by which the income is produced.

Discount and the laws of price. It is essential that we understand exactly how this process of discount is related to the general laws of price which we have previously developed. We have found that competitive prices are determined either by demand and supply, or by cost of production, or (occasionally) by the will of a sole buyer. We have seen that the fundamental thing back of demand is the marginal utility of the good in question. Marginal utility is what governs the action of the buyers, and it is also the controlling influence upon sellers in the case of those goods which are not reproducible, such as land, and whose exchange therefore involves sacrifice of use by the seller. Even in the case of monopoly prices demand plays an important rôle, and here also marginal utility is the controlling factor.

Now the marginal utility of an instrument of capital to any person is based upon his estimate of the utility of the income which he anticipates from it. And since all such income is in the future, the discount process is necessarily involved in arriving at its present value, upon which depends the utility of the capital instrument.

This mechanism appears most clearly in the case of those capital instruments whose value is in no way related to cost of production, either because they have no cost of production or because they are not reproducible. Land is an important example of this type of capital. The value of a given acre of land will be set by the forces of demand and supply in the real estate market. Both buyers and sellers will base their offers upon their individual estimates of the value of the services which may be anticipated from this acre. If the consensus of judgment is that land of this sort may be expected to return a future net income of $10 a year forever, and if the interest rate is five per cent, this plot of ground will have a sale price of $200; *i.e.*, $10 \div \frac{5}{100}$.[1] A rise in the expected value of the net income from $10 to $20 per year, the interest rate remaining the same, will double the value of the acre; a fall in the value of the income, the interest rate again remaining the same, will diminish the value of the acre. Obviously this is true of all lands. The same may be said of all goods which, though produced by man, cannot be reproduced. There can be, for example, but one Panama Canal, and one Suez Canal, one system of piers in

---

[1] The general formula for the present worth of a uniform perpetual income is $P = i \div r$, in which P stands for the present worth, i stands for the periodical payment, and r stands for the rate of interest.

New York Harbor, one power dam on the site of a given waterfall. If a company were organized to buy the Panama Canal and operate it for profit, the price the stockholders would be willing to pay would be determined, through the discount process, by their expectation of the net yield to be derived from toll charges. The fact that it cost, let us say, two hundred millions to build the canal would not figure in their computation of its present value. So it is with all capital instruments which, though they have a cost of production, cannot be duplicated.

The relation between future income and present value appears very clearly also in the case of reproducible goods when the income realized turns out to be less than was anticipated. Buildings, we know, are reproducible goods, made by groups of men to be sold at a profit. The construction of dwellings, office buildings, hotels, and factories will not be undertaken unless the builders expect the selling price to be as high as the cost of production. But after the costs involved in the erection of a given building have been borne and the building is offered for sale, the price will be determined by the buyers' estimate of the present worth of its future products. A summer resort constructed at great expense in an unpopular place will command a price lower than its cost of production because the discounted value of the net income to be derived from its operation is low. For the same reason an office building erected in a decaying town will have a value below its cost of production; so also will a railroad which runs through deserted territory, or a machine whose products cannot be sold, or a department store which has lost its patronage, or any other existing capital instrument whose income, when discounted, gives it a low present value.

Discounted income and cost of production. Now there is a vast number of goods which are staples in the market. They are constantly being reproduced as a matter of routine business practice; day after day they appear in a steady stream upon the markets of the world. The producers of these goods, after they have offered their supplies for sale, must be content with a price determined by the laws of market price, in the operation of which the discount process will govern the demand of the buyers. But what if the estimate thus placed by the buyers upon any given quantity of the goods differs from the cost of producing this quantity? The discrepancy will be resolved, as we have learned, through the power of the producers, on the one hand, to avoid future costs by giving up the production of these goods or reducing their output or, on the other hand, to increase future costs by increasing their output. Their decision on this point will, after an interval of time, exert an influence upon the value of the capital goods they are producing. If they produce an increased quantity of the capital goods, the benefit derived from them will increase in quantity and command a lower price; hence the discounted value of the capital goods will fall. On the other hand if the producers reduce their output of the capital goods, the peculiar benefits which flow from these goods will become scarcer and their price will rise, hence raising the discounted value of the capital goods. In other words, it is the willingness or unwillingness of the producers to bear future costs which, by affecting the value of the income from capital, brings marginal utility and cost of production into harmony and so influences the selling prices of capital instruments.

In the case of standard and staple goods, the markets for which have settled to a point of relatively stable equilibrium, the producers may confidently take the present price as an indication of what the future price will be. If the present price of the good stands very much above the cost of making it, existing producers will be tempted to expand their output; while at the same time other enterprisers will be tempted by the prospect of large profit to engage in its production. These activities will cause the discounted value of the instrument to fall as its income grows cheaper, due to increased plenty.

If the industry is subject to the condition of increasing costs the marginal cost of producing the instrument will rise as attempts are made to supply it in larger quantities. As a result of both of these tendencies, the discounted value and the marginal cost, which at the outset were different from each other, will begin to draw together. On the other hand if the present price of the capital instrument had been below the marginal cost, enterprisers desirous of avoiding the losses caused by this situation would abandon the production of the instrument entirely or contract their output of it. This restriction in the supply of the instrument would cause the income it produces to grow scarcer and higher in price, with the result that the discounted value of the instrument would rise. And the diminution in the production of the capital instrument would reduce the marginal cost of production. Again there would be a tendency for the two — cost of production and discounted value — to approximate each other.

If, on the other hand, production is subject to the condition of constant costs, as the reader will recall is the case in many important industries, price will be determined solely by the cost of production. But the discounting process will be no less potent in its influence upon the demand of the buyers, and producers will, as described in the preceding paragraphs, by either reducing or increasing their output determine such quantity of production as will cause the marginal utility to agree with the cost.

EXERCISES

1. Suppose $100 to be deposited in a savings bank which pays four per cent interest on such deposits.
    (*a*) What would this deposit be worth in three years' time if no money were withdrawn in the meantime?
    (*b*) What would it be worth in six years?
    (*c*) What would it be worth in fifteen years?
    NOTE. Solve (*a*) exactly. For (*b*) and (*c*) work out the formula only.

2. Suppose the interest rate to be five per cent. What is the present value of a promissory note for $1,000 due in two years?

3. Suppose the interest rate to be five per cent. A bond is issued which is described as follows: "$1,000, ten year, six per cent bond." This description means that the owner of the bond will receive $60 per year (six per cent of $1,000) for ten years and a redemption payment of $1,000 at the end of the tenth year. What is the present value of this bond? Work out the formula only.

4. A business man obtains a contract from which he draws a net profit of $5,000 per year, the contract expiring in six years. At what price could he sell this contract at the end of the fourth year, if the interest rate at that time is four per cent?

5. A certain piece of land can be made to return a perpetual net income of $500 a year. What is its present value, (*a*) at a five per cent rate of interest? (*b*) at a four per cent rate? (*c*) at a six per cent rate?

6. In general, if income value remains unchanged what is the effect on the value of capital, (*a*) of a rise of the interest rate? (*b*) of a fall of the interest rate?

Suggestions for Further Reading on Part II, the Laws of Price

Sections devoted to the subjects of price and value will be found in all of the general economics texts listed at the end of Part I. The following will be found especially useful:

FISHER, I. *Elementary Principles of Economics.* New York, 1912
MARSHALL, A. *Principles of Economics.* Seventh edition. London, 1916
TAYLOR, F. B. *Principles of Economics.* New York, 1921
SHEARMAN, H. P. *Practical Economics.* New York, 1922

Among the works devoted more particularly to the subject of price and value, the following are suggested:

HENDERSON, H. D. *Supply and Demand.* New York, 1922
WIESER, F. VON *Natural Value.* (English translation) London, 1893
DAVENPORT, H. J. *The Economics of Enterprise.* New York, 1913
MEEKER, J. E. *The Work of the Stock Exchange.* New York, 1922

See also the books on money and prices listed at the end of Part III.

5600
.04
200.00

10000
05)500000

PART III

MONEY, BANKING, AND EXCHANGE

$10000
.05 | 500.00

CHAPTER XX

MONEY

The modern economic structure is reared on the foundation of division of labor and exchange of products. This was the fundamental thesis of Part I of the present book, and Part II was devoted to the principles of value, which regulate exchange and through exchange control the production and distribution of wealth and services. Exchange, in its modern omnipresence and complexity, would be impossible without an efficient medium, which brings us to the subject of money, the medium of exchange. Money and its correlative, banking, furnish therefore the central theme of the third part of this book.

Examples of primitive money. Money appears today and has appeared in the past in such a variety of forms that one's first steps in its study are apt to produce a sense of bewilderment. There are coins of gold, silver, nickel, bronze, and other metals. There are pieces of paper inscribed with promises and other contracts. Cattle and sheep have been used as money, as well as wheat, corn, tobacco, musket balls, beaver skins, fish, meat, shell beads, gold dust, and so on indefinitely. Monetary history discloses that each community has inclined to adopt as its money some commodity or commodities of common use among the people, something that is desired by most of the people most of the time.

One of the earliest forms of money of which we have knowledge is the sheep and cattle of peoples on the pastoral stage of economic development. Most of the people derived their living from their flocks, and these represented their principal form of wealth. One would always welcome an addition to his flock, and thus it became the natural thing to use the domestic animals as money. The word "pecuniary," derived from the Latin word *pecus* for cattle, is an interesting heritage of the early stage in which cattle and money were in a sense synonymous.

The early settlers of Virginia and the other southern colonies found the cultivation of tobacco their most profitable occupation. A considerable part of the people was engaged directly or indirectly in the tobacco business, either as planters, carriers, dealers, or exporters. So tobacco was generally acceptable; even the individual who neither smoked nor chewed accepted tobacco readily, since he knew he could always exchange it for other things.

The American Indians made, out of the inner whorls of certain periwinkle shells found along the seacoast, small beads, which they prized highly as ornaments. They were strung on thongs and braided and woven to make necklaces, bracelets, belts, and sashes, and the wealth and social standing of the brave was apt to be indicated by the quantity of this "wampum" which he could display. Being thus universally desired, wampum became the money of the Indians. So far we have merely another example of the ordinary historical origin of money. But note what happened when the white settlers came and opened commercial relations with the Indians. An important trade arose in skins and other forest products, particularly beaver skins, which the colonists could always export to Europe on favorable terms. Since an Indian would always accept wampum in exchange for beaver, the white traders were themselves perfectly willing to accept it. And since the traders would always accept it, other whites who had dealings with them were willing to receive it. Before long any white man would receive wampum, knowing that he could always use it for making purchases, not only from the Indians but from the other whites as well. So there grew up a form of money of no intrinsic value to the people themselves, but drawing its value entirely from its exchangeability, acceptable to each because acceptable to all, and based finally upon its acceptability by a race of barbarian neighbors.

The common characteristic. From these few examples one can derive a mental picture of the origin of money and of the almost infinite number of commodities which have served as money in various places and at various times. But, what is of more significance, one begins to sense a common feature among all this otherwise bewildering variety of monetary expedients, a feature which is at first likely not to be appreciated because of its very obviousness. This is the character of general acceptability in exchange for other things. Some of the early forms of money were clearly desirable for their own sake. Indeed all money doubtless arose out of articles thus desirable. But others, such as the wampum used by the New England colonists, were of no use whatever for the direct satisfaction of wants. Such forms of money were desired only because of their general acceptability in exchange.

It is the fact that money is acceptable generally which makes it acceptable to each individual, if we may be permitted a statement which appears to involve reasoning in a circle. If the person who has valuable goods to sell is willing to accept in exchange a coin of gold, or a piece of paper engraved with the promissory note of a bank, or a bale of tobacco leaves, it is not ordinarily because he has any need of these particular articles; they may be quite incapable of satisfying directly any of his wants. He accepts them simply because he knows that at any time he chooses he will find other people willing to accept them in exchange for whatever they may have to sell to him. This is the fundamental characteristic of money; so long as this qualification is met there may be the utmost diversity as to other features.

Money is local. The general acceptability of money is always more or less local. Most American merchants would refuse to accept a Bank of England note in payment for goods sold, and a Canadian quarter is not generally acceptable in America except close to the border. The foreign traveller has to exchange his own money for the moneys of the particular countries which he proposes to visit. In general each nation has its own monetary system, and the money of one nation is usually not acceptable within the borders of other nations. By general acceptability we mean acceptability within the particular community.

Definition of money. We have now all the essentials of the scientific definition of money, which was stated but not explained in a previous chapter. *Money consists of articles of wealth and property rights which are generally accepted in a certain community in exchange for other wealth and services.*

Functions of money. The principal function of money is to facilitate exchange, or in other words to serve as a "medium of exchange." Money thus enables people to escape from the clumsy and ineffective device of barter. As a corollary to its main function, money furnishes the unit of value, that universal unit by which alone all kinds of wealth and services may be measured and totalled. As another corollary, money enables us to break up and separate our sales and our purchases according to our convenience, something which is accomplished only with the greatest difficulty under the régime of barter. These are simple statements. Yet they involve principles which are at the very foundation of all modern business and indeed of virtually all economic activity of civilized mankind.

Gold and silver. At various times and places a wide variety of commodities has been employed as money. Yet as time has gone on most of these monetary expedients have been gradually eliminated, till in modern times the civilized world has quite generally settled down to use of the precious metals, gold and silver, and of certain property rights based more or less directly upon these precious metals. The explanation of this important development will appear if we inquire what are the peculiar qualities of gold and silver which make them especially fitted to serve as money.

Qualities of the ideal money commodity. First of all, gold and silver have always and everywhere possessed in unusual degree the fundamental quality necessary to any monetary commodity, namely *general desirability.* The precious metals have always been desired for the fabrication of ornaments, and in modern times they have also important mechanical uses, as for example in dentistry. It would be difficult to think of any other commodity more universally desired.

In the second place gold and silver have *great value in small bulk.* This is an important consideration, since it makes possible the carrying of a considerable value without being burdened by the weight. If one considers other commodities of great usefulness which might conceivably serve as money, such as wheat, coal, salt, etc., he will recognize at once the inconveniences and absurdities which would follow the use of such a bulky medium of exchange. This quality of the precious metals also facilitates the transport of money easily and cheaply from one place to another, locally and internationally.

Again gold and silver have the advantage of *durability*. Age and exposure do not materially affect them, and they may thus be kept indefinitely without danger of loss from physical deterioration. This is of course an important matter, if money is to perform its function of enabling people to make their sales at one time and their purchases at a later time; that is, to hold in the form of money their command over goods and services generally. No other commodity which might conceivably serve as money possesses so completely this quality of durability, while most are patently deficient. The Virginia colonists for example were always suffering loss through the deterioration of their stores of tobacco money.

Another desideratum of a convenient money commodity is *uniform quality*. All the possible advantages of monetary exchange are not obtained when each piece of money must be examined and tested as to its quality. Wheat would make a poor kind of money for this reason; so would coal or cotton or indeed almost any other possible commodity. The Virginia legislature had to take cognizance of the distinction between "good tobacco" and "bad tobacco," and it was necessary to have officials to examine and grade the tobacco. Even so, the colonist had always to be on his guard against receiving payment in inferior or spoiled money. This was also a serious defect of the wampum currency as used by the New England colonists. The precious metals, on the other hand, have perfect uniformity. When to this is added the quality of durability, which we have just noted, it results that an ounce of fine gold is exactly like any other ounce of fine gold, whether it came from a California mine in 1849 or was mined last year in South Africa or the Klondike.

The convenience of a monetary commodity is greatly enhanced by *divisibility,* which facilitates the making of payments of all possible amounts. Many commodities, such as coal, wheat, cotton, etc., possess this quality; others, such as cattle, do not. Gold and silver may be divided and made into coins of any desired size without loss of value. Gold and silver have also the obvious advantage of being easily recognized, of *cognizability.* They are not easily confused with other substances, and counterfeiting is difficult.

Stability of value. Finally gold and silver have, relatively at any rate, stability of value. This, like physical durability, is essential in order that wealth may be held in the form of money and still more essential in connection with the making of contracts for future payments, loans, deposits at the bank, and investments in general. Most commodities fluctuate considerably in value from day to day and may undergo great changes in a period of months or of years. Such changes in the value of money, in which are measured the values of all other goods and services and in whose terms all kinds of contracts are drawn, is extremely embarrassing and the cause of unjust loss on the one hand and undeserved gain on the other. They produce the same sort of confusion and injustice as would result from the use of a yardstick of varying length or a quart of fluctuating capacity. Lack of stable value was the most serious of the many defects of the Virginia tobacco currency and the rock upon which that particular monetary system finally went to pieces. The price of tobacco fell from three shillings six pence per pound in 1628 to only a penny in 1665, with disastrous results, which the reader can easily imagine and which will be more fully appreciated after our study in a later chapter of the relation between money and prices.

Now gold and silver are by no means perfect in this respect; they are subject to very material fluctuations in value, a defect which has led to serious consideration of plans for an entirely different sort of monetary system not based on the value of any one commodity. Putting aside for the present the study of such plans, we note here that, of all the commodities which might serve as money, gold and silver, or at least gold, possess in the greatest degree the quality of stable value. This is in large part due to their durability. Gold and silver are not generally destroyed by use, as is the case of most other commodities, such as wheat, cotton, coal, salt, etc. The world's stock of these metals is thus a slowly increasing mass which at any given time contains much of the production of past years even back to prehistoric times. The effect upon this total stock of one year's production, whether large or small, is not very important and has smaller influence upon the supply and the value than in the case of another commodity, like wheat, where the stock at any time is generally limited to about one year's production. Yet even this peculiarity of the precious metals has not prevented great changes in their value, as we shall later see. The stability of gold and silver is after all only a relative stability as compared with other commodities.

Superiority of gold and silver. We have now passed in review a catalogue of some seven qualities required of any commodity which is to perform perfectly the monetary functions. There is no commodity in the world which entirely meets these seven requirements. Gold and silver have, through the long test of the world's history, proved themselves most nearly qualified, which explains the fact that they have gradually displaced virtually all other commodities in the monetary systems of the world. Gold has in fact proved itself better qualified than silver and, since the experiment of using both gold and silver on equal terms has proved a failure, the modern nations have finally settled upon gold alone as the basis of their monetary systems, using silver and other metals only in a subsidiary capacity.[1]

Government control of money. While the historical data are not complete nor entirely unambiguous, it is safe to say that money originated in each community naturally and unconsciously as the result of slowly established custom. That is, there was not at the start an inquiry and agreement upon the commodity which should serve as money. People simply fell into the habit of accepting some one commodity in preference to others in exchange for goods

[1] China is the only nation of importance now using the silver standard.

and services which they had to sell, much as boys at a certain stage of their development fall into the habit of using marbles as a medium of exchange. Thus the monetary system became firmly established in the mores of the community, and eventually the government took cognizance of it and proceeded to define and enforce it. For centuries the control of the monetary system has been jealously guarded by kings and rulers as one of the most precious attributes of sovereignty. Today the system in all its details is determined by law, and it is this modern legal monetary system with which we are concerned. Government (1) specifies exactly what kinds of wealth and property shall be money, fixes the monetary unit, defines the different kinds of money and their relations to each other, (2) coins the metallic money and engraves and prints the paper money, and (3) makes and enforces rules regarding legal tender. We may now proceed to examine the essential features of the monetary system as it prevails among the civilized nations of today, not attempting to describe the system of any particular country, but taking our illustrations principally from the monetary system of the United States.

Money includes coins of gold, silver, and certain other metals or compositions, such as copper, nickel, bronze, etc., and paper money. The latter is generally of two forms, either certificates representing the right of the bearer to coin held by the government, or promissory notes either of the government or of a bank or banks, promising to pay coin or some other kind of money to the bearer. We shall later study each of these kinds of money in more detail.

The monetary unit. The law prescribes the monetary unit in which the various kinds of money are stated and measured. In the United States the unit is the dollar, consisting of 25.8 grains of standard gold, standard gold being a mixture of nine tenths pure gold and one tenth an alloy of copper. The dollar thus contains 23.22 grains of pure gold and 2.58 grains of copper. Similarly the British monetary unit is the sovereign, or pound, containing 123.274 grains of standard gold; British standard gold being eleven twelfths fine, the sovereign contains 113.001 grains of pure gold and 10.273 grains of alloy. Each of the other nations has its own monetary unit, the franc in France and Switzerland, the lira in Italy, the mark in Germany, etc. By agreement, France, Belgium, Switzerland, Italy, and several others of the European nations have adopted the same unit, though it goes by different names in the several countries. Note that the unit is always a certain quantity of metal, whose weight and composition is exactly prescribed by law. The unit may or may not be represented by a coin. There is for example no gold dollar in the United States, since the coin would be too small for convenient use.

Coinage. The purpose of coinage is to make money which the people may use safely and conveniently. In the early days in California, after the great gold discoveries of 1849, gold became the medium of exchange by exactly the same natural process which led to the cattle money of the ancient pastoral tribe. Quantities of gold dust or nuggets were measured out in making payments. This was obviously inaccurate and inconvenient; there was also the opportunity for misunderstanding and fraud. When gold is coined, everyone knows exactly the weight and composition of the coin he pays or receives.

Coinage might conceivably be left to develop as a private enterprise. This actually occurred in California. Business men and corporations made coins of various sizes, compositions, and designs, and the people found these coins much safer and more convenient than the gold dust money. However the most perfect coinage is not obtained until the government takes the business over and puts a stop to private coinage. Then there becomes one unit, uniformity takes the place of variety, and the suspicion and uncertainty that might attach to coins issued even by well-known business corporations is replaced by absolute confidence. Hence the logical reason for government monopoly of coinage, whose development it must be confessed was aided also by the cupidity of many an early monarch.

Free coinage. The government obtains the metal for coinage in one of two ways; either it buys the metal in the open market or it allows *free coinage.* The former is simple enough, but the latter requires some explanation. Free coinage may be defined as *that system by which the government is legally required to coin for any person any amount of a particular standard metal which he may bring to the mint.* For example there is in the United States free coinage of gold. Any person may bring to the mint any amount of gold of the standard prescribed by our coinage laws, and the mint will give back to him the same weight of gold coin. Some governments perform this service for nothing; others make a charge sufficient to cover the cost of the coinage. In the United States the practice has varied. At first there was no charge; from 1853 to 1873 the charge was ½ of one per cent; from 1873 to 1875 it was fixed at ⅕ of one per cent; since 1875 there has been no charge. When free coinage is granted without a charge for the cost of the operation, the term *gratuitous coinage* is employed. Note that there is free coinage, as defined above, whether the coinage is gratuitous or not.

The metal brought to the mint for coinage is often not standard; that is, not the exact mixture of pure gold and alloy prescribed by the law. In that case the mint or the government assay office will assay and refine it, charging the cost to the one who brought the gold or ore. The gold may then be returned to its owner or made into coin for him, and in the latter case there is a further charge for the alloy. In practice in the United States, the owner usually does not either take the gold or wait for its coinage, but takes payment in lawful money; that is, sells the gold to the mint. The mint then holds the gold either in "fine bars" for commercial use or in "mint bars" for coinage at its discretion.

Making and regulation of paper money. Coinage is the making of coins or metallic money. Closely related is the making and regulation of paper money. Whatever the nature of the paper money, it has been found desirable to have the government prescribe its form and control its issue. This is for much the same reasons as dictate the government monopoly of coinage; *i.e.*, to prevent misunderstanding, counterfeiting, and other frauds, and to obtain the obvious advantages of uniformity. Thus in the United States all kinds of paper money, even the notes issued by the national banks, which are privately owned corporations, are designed, engraved, and printed by the national government. Great care is exercised, by the choice of a peculiar kind of paper, by intricate designs, and by careful engraving and press work, to give a distinctive appearance and make counterfeiting difficult. The government exercises regulation over the paper money in various other ways, which will be noted in connection with our later study of the several kinds of money.

Legal tender. Every monetary system includes various kinds of money, all of which are " generally accepted " in exchange for anything else and in settlement of all contracts which call for money payments. Yet the legal obligation of the creditor to accept payment is not the same with respect of the several kinds of money. For example if a promissory note falls due anywhere in the United States the debtor may offer gold coin in payment and the creditor will be compelled to accept it; he cannot insist upon payment in some other kind of money for which he may have a preference. On the other hand if the debtor should offer national bank notes or silver certificates the creditor might refuse to accept them and compel the debtor to make payment in one of certain other kinds of money. This brings us to the subject of *legal tender*, about which there is a good deal of popular misunderstanding. Legal tender is defined as *any kind of money which according to law must be accepted when offered in payment of any obligation expressed in terms of the country's monetary unit.* Examples have already been given; thus in the United States gold coin and gold certificates are legal tender, national bank notes and silver certificates are not. This is a matter of legal enactment.

It must be clearly understood that the legal tender law has nothing to do with contracts or other obligations not payable in money. Neither does it apply to contracts which by their terms are payable in a particular kind of money. There is nothing to prevent people making contracts payable in any kind of money they may agree upon. A man might agree with his employer to work for wages of twenty-five dollars a week payable in national bank notes or in silver half dollars or in any other kind of money. The legal tender law would not enable the employer to avoid payment in the specified kind of money, by offering gold coin for example. In other words the legal tender law does nothing to abridge the right of making contracts or the enforcement of contracts when made. Only when the agreement is in terms of dollars (or other monetary unit), no particular kind of dollars being specified, does the legal tender law apply. It then serves, in case of dispute, to settle just what is meant by the word "dollar." Since practically all money agreements are thus generally stated or understood, the legal tender law is actually of wide application. Its normal purpose is (1) to protect the creditor from being forced to accept anything but the best kind of money when different kinds are in circulation and (2) to protect the debtor against a capricious creditor who might refuse to accept a perfectly good kind of money or insist upon receiving some form which the debtor could not conveniently provide.

Classification of money. Before embarking upon the investigation of the separate species of money, we shall find it desirable to have a general notion of the monetary system as a whole and of the principal parts of which it is ordinarily composed. As a first step, the following classification[1] will prove useful:

1. Standard money.
2. Fiduciary money.
   (*a*) Representative money.
   (*b*) Token money.
   (*c*) Credit money.
3. Fiat money.

As the name implies, *standard money* is the basis of the normal monetary system. It is that money whose value depends upon the value of the material of which it is made; *i.e.*, it has "intrinsic" value. Also it is the money which gives value to the various kinds of fiduciary money in the second group.

Those forms which depend for their value, not upon their intrinsic worth, but upon their relation to the standard money are *fiduciary money.* There are three important classes. (*a*) *Representative money* consists of certificates testifying to the fact that the government holds standard money which will be delivered to the bearer of the certificate on demand. Representative money is thus of the nature of a warehouse receipt. (*b*) *Token money* consists of coins containing less than their nominal value of metal and redeemable in standard money. The token coins are usually of smaller denominations than the monetary unit. (*c*) *Credit money* consists of notes in which either the government or certain banks agree to pay standard or some other kind of money to the bearer; *i.e.*, credit money has the nature of a promissory note.

---

[1] Classification must not be taken too seriously. It is never to be regarded as an end in itself; its purpose is merely to facilitate investigation. Classification therefore may always be more or less arbitrary, and it is not contended that the classification of money here presented is the only one that might be set up and defended. It does commend itself to us as on the whole the most useful for the present purpose.

In the case of *fiat money* we part company from the standard money altogether. Fiat money has no intrinsic value itself and is not redeemable in any other kind of money. It is money, ordinarily paper, simply because the government makes it so, hence "fiat," and its value depends on its usefulness as money and its quantity.

With this brief general analysis as a background, we may now proceed to study somewhat intensively these several kinds of money, commencing with a number of important problems connected with standard money.

The value of standard money. In the United States, as in most other countries today, the standard money is gold coin. The value of a gold coin is due to the value of the gold in it, not to the government stamp upon it. A gold eagle might be hammered or melted down till it was a mere lump of gold; yet that lump of gold would still be worth ten dollars. If taken to the mint it could be coined into an eagle, but would gain nothing in value thereby. A gold eagle contains exactly 258 grains of standard gold, and a lump of standard gold weighing 258 grains tends always to be worth exactly ten dollars.

This relation arises from (1) free coinage of gold, (2) freedom to melt down gold coin, and (3) the ordinary laws of price. Only a brief explanation will be required to make this clear. Let us suppose that at a certain time the equilibrium were disturbed and the gold eagle were worth more than a lump of standard gold bullion weighing 258 grains. Those having gold bullion would seek to exchange it for gold coin, at the rate of 258 grains for a ten dollar coin; they might even be willing to offer a slight premium for coin, that is, to offer more than 258 grains, if they had to. But they would not have to, since anyone can take 258 grains of standard bullion to the United States Mint and receive a gold eagle without charge. Hence there could be no premium on gold coin. People would simply take bullion to the mint and have it made into coin. As they did so, the quantity of bullion would be decreased, which would, according to the laws of price, tend to decrease its supply on the market and so to raise its value. Similarly the additions to the stock of coin would increase its supply and tend to lower its value. The discrepancy between the value of coin and bullion, which was our assumption, would thus be reduced from both sides till finally equilibrium was restored and the values of coin and bullion were identical.

Consider now the opposite case, a situation in which there was a special desire for bullion, making 258 grains of gold bullion worth more than ten dollars. People might be willing to offer more than ten dollars for 258 grains of gold, if they had to. But of course no one who had gold coin would actually make such offer, since by the simple process of melting anyone can turn a gold eagle into 258 grains of bullion without loss. As this was done, gold coin would become scarcer, tending to raise its value, and gold bullion would become more abundant, tending to lower its value. Ultimately the two values would have to become the same.

Clearly the condition in which gold coin is worth either more or less than the value of the gold of which it is composed is one of unstable equilibrium, which tends immediately to be corrected. This demonstrates the general principle that the value of standard money tends always to be equal to the value of the substance of which it is made; *i.e.*, to its intrinsic value.

Essentials of standard money. It will be inferred, and correctly, that standard money is always an article of wealth, a commodity, not a property right. It should also be noted that we have assumed free coinage, which is always an attribute of the standard money. Standard money is likewise always legal tender, if there is any legal tender law at all. Being the best money there is and of full intrinsic value, everyone is permitted to pay his monetary debts in it. It would be most awkward for the debtor if his creditor could refuse the standard money and insist upon payment in some other kind of money, and it is surely no hardship to the creditor to compel him to accept the best money there is.

The gold standard. Gold is today the monetary standard of the world.[1] In times past silver was frequently so used, and it is only recently that certain important countries, such as Mexico and India, have abandoned the silver standard in favor of gold or gold exchange. China alone clings to the silver standard. Gold gradually gained favor over silver because centuries of experience showed that on the whole it better met the requirements which we have noted. When the majority of the leading commercial nations had adopted the gold standard, the others were hastened into line by the further advantage of uniformity. No economic principles are involved in the silver standard which are different from those affecting the gold standard.

Bimetallism. There is still a third standard which, though no longer existing, has played a very important rôle in monetary history and involves economic principles of great interest and importance. This is the double standard, or bimetallism, in contrast to which the other standards are called respectively gold monometallism and silver monometallism. In a bimetallic system both gold and silver are concurrently used as materials for the standard money, and the monetary unit is defined as either a certain weight of standard gold or a certain weight of standard silver, the law prescribing also the exact composition of each standard metal. The ratio between the weight of pure silver in the silver unit and the weight of pure gold in the gold unit is called the coinage ratio, mint ratio, or legal ratio. There is free coinage of both gold and silver, and the standard coins of either metal are unlimited legal tender, these being always the two essential attributes of standard money.

---

[1] There are, it is true, certain countries, notably India and the Philippines, where the standard is not gold coin but exchange upon some other country having the gold standard. This situation in no way conflicts with the monetary principles which we have developed. The discussion of the "gold exchange standard" is a matter of technical detail into which it is not necessary for us to enter.

An example. The United States had bimetallism from the establishment of the national monetary system in 1792 down to 1873. The monetary unit was the dollar, which (after slight changes in 1834 and 1837) was defined either as 25.8 grains of standard gold or 412.5 grains of standard silver. Standard metal contains (since 1837) nine tenths of pure gold or pure silver and one tenth of copper. The gold dollar therefore contained 23.22 grains of pure gold, and the silver dollar, 371.25 grains of pure silver. The coinage ratio was $\frac{371.25}{23.22}$ or $\frac{15.988}{1}$, being almost exactly sixteen to one. There was free coinage of both metals (except when the coinage of silver dollars was stopped for a time after 1804), and the coins of both metals were unlimited legal tender. There was therefore a real system of bimetallism.

Problems of bimetallism. In connection with bimetallism there arise certain important questions in addition to those which relate to monometallism. If debts may be paid either in gold coin or in silver coin and if either kind of coin may be freely obtained at the mint in exchange for the respective bullion, how will the people act? Will it make any difference which metal is used, and if so what causes will determine the choice? Will both metals be actually used? These questions present an interesting theoretical problem, they have been the subject of violent political controversy in the United States, and they give special interest to research into monetary history.

Mint ratio and market ratio. The theoretical analysis is fairly simple. Its starting point is consideration of the relation between the mint ratio and the market ratio, the latter being the ratio between the value of a given weight of gold and the value of the same weight of silver on the market. This ratio is of course the result of the respective values of gold and silver; these values are determined in their turn by the demand and supply in each case;

and finally demand and supply depend on the one hand upon the uses of the metals for all purposes (not merely for money) and on the other hand upon the conditions of mining. This ratio can obviously not be controlled by law and it is subject to frequent change, as is the ratio between the values of any two commodities.

When the two ratios are equal. At any given time the market ratio must be either equal to, or less than, or greater than the mint ratio. Taking the first case, let us assume that a bimetallic system has a ratio of 16 to 1 and that the market ratio between silver and gold is also 16 to 1. A given amount of gold will exchange on the market for 16 times its weight of silver, and a given amount of gold will make at the mint exactly the same value in coin as 16 times its weight of silver. There will be no reason for taking one metal to the mint rather than the other; both metals will be coined.

When the market ratio is less than the mint ratio. Now let it be supposed that, the mint ratio still being 16 to 1, the market ratio is something less, say 15 to 1. Now anyone having silver bullion and desiring coin would find it to his advantage to exchange his silver for gold before going to the mint. For 15 ounces of silver he can buy 1 ounce of gold, which at the mint will give him just as many dollars as if he had taken there 16 ounces of silver. Therefore no one will take silver to the mint; only gold will be coined. Furthermore there will similarly be advantage in melting down and using as bullion any silver coins which may then be in circulation. Silver coins weighing 15 ounces can be exchanged for an ounce of gold, which at the mint will be made into coin of the same face value as 16 ounces of silver coins; the exchange has netted 1 ounce of silver. Silver coins thus tend to disappear from circulation; gold is the only metal being coined, and the only coins in circulation are gold. The effect is practically the same as gold monometallism.

When the market ratio is the greater. The opposite results follow in case the market ratio is greater than the legal ratio. Suppose the legal ratio is 16 to 1 while the market ratio is 17 to 1. Anyone having an ounce of gold can now exchange it for 17 ounces of silver, 16 of which taken to the mint will yield him the same number of dollars as would have been obtained by having his ounce of gold coined. There is also gain to be obtained by melting down gold coin for bullion. Therefore only silver will be coined, and the gold coins already in circulation will tend to disappear. The effect is practically the same as silver monometallism.

Gresham's law. It will be noted that it is the metal which the mint overvalues in comparison with the market ratio which drives the other out of circulation. Coins of the first metal may be obtained from the mint more cheaply than coins of the other, for the reason that the mint places a higher value on that metal in terms of the other than does the market. This principle is part of a broader generalization, known as "Gresham's law," which is fundamental to all monetary theory and may be stated as follows: *When two or more kinds of money of unequal value are in concurrent circulation, each being available for payments, the inferior tends to drive the better out of circulation.* This law may be seen in operation whenever the conditions are fulfilled. It applies when anyone, selecting from a pocket full of bills, makes his payment with the worn and soiled bills and keeps those that are new and crisp, as well as when new heavy coins are driven out of circulation by those that are worn or clipped, and when the cheaper coins in the bimetallic system drive out the dearer. The better money may be hidden away and hoarded, or exported for payments abroad, or melted down for bullion. The inferior remains in circulation. We shall rub elbows with Gresham's law more than once in the following pages.

The operation of Gresham's law takes time. The reader scarcely needs the reminder that these results of a discrepancy between the two ratios will not come instantaneously. All persons concerned are not fully aware of these principles, the facts of the situation may not be recognized immediately, and there may be some who are not absolutely keen to take advantage of their opportunity. Since however the actual business of taking metal to the mints and of using metal in the arts is in the hands of a few experts, this qualification is of little moment as regards the choice of the metal which shall be coined. The disappearance from circulation of the metal which is undervalued at the mint may take more time, though even this result will not be long delayed, particularly in a country like the United States where very little of the standard coin is ever actually used as pocket money. When most of the standard money is in the reserves of the banks and the government's vaults, it is certain that expert knowledge will quickly take full advantage of the situation.

Will a discrepancy between the ratios correct itself? Before reaching our final conclusion we have also to observe that the forces set in operation by a discrepancy between the mint ratio and the market ratio have themselves a tendency to bring the ratios back to equality. Take the case of a market ratio less than the mint ratio between silver and gold. The existing silver coins begin to pass out of circulation, and new coins are made only of gold. The whole task of furnishing standard money, formerly shared by gold and silver, is now about to be placed upon gold. This means an increased demand for gold and a consequent tendency for its market value to rise. On the other hand, silver being no longer used for standard money, the demand for silver is less, with a tendency to a decline in its value. The rise in the value of gold and the decline in the value of silver tend of course to increase the market ratio between them and so to correct the original discrepancy. It is possible theoretically that the equilibrium between the market ratio and the mint ratio might thus be restored before all the standard silver coins had been eliminated. From that point the coinage of silver would be resumed, and the bimetallic system would go on as before, except that some of the silver coins would be definitely lost and the circulating medium would contain relatively more of gold and less of silver than previously. The reader can readily enough work out for himself the opposite train of events that would be set in motion if the market ratio were greater than the mint ratio.

The theoretical question. Of the truth and importance of this principle there cannot be the slightest question. Its potency to perpetuate effective bimetallism in the face of fluctuation in the market ratio of gold and silver has however in the heat of political controversy, been greatly exaggerated. At the time when the restoration of the double standard by the United States (abandoned in 1873) was being vigorously demanded, some of the extreme friends of bimetallism believed that, through the operation of this principle, bimetallism would always work and neither metal would ever be entirely driven out of circulation, no matter what the market ratio might, except for bimetallism, have become. On the other side it was claimed that the automatic corrective effect of bimetallism was so slight as to be powerless to restore the market ratio if it ever departed from the mint ratio by more than a trifling amount. The theoretical argument, during the presidential campaigns of 1896 and 1900, was fierce and stubborn.

The practical answer. Now it is probably not possible to find a precise quantitative answer to this problem by the methods of theory alone, and much of the wordy argument of 1896 and 1900 appears as pure waste, particularly when it is noted that the disputants might have found entirely adequate answer to their question in the past monetary experience of the United States, not to mention foreign countries. For this purpose, as well as for verification of all our principles of the monetary standard, let us now glance at the monetary history of the United States, with which the next chapter begins.

CHAPTER XXI

MONEY (Continued)

Bimetallism in the United States. When the American people started their career as an independent nation, their money was a make-shift collection of paper notes and foreign coins left over from the colonial and Revolutionary periods. The greater part consisted of notes of the Continental Congress, of the several colonies or states, and of a few banks. Except for some silver shillings and sixpence previously coined by the Colony of Massachusetts Bay, there were no American coins. The most common coin was the Spanish peso or dollar. A national monetary system was first established by the mint act of 1792, only three years after the formation of the government of the United States. The unit was the dollar, of either gold or silver, though there was no one dollar gold coin, the gold coins being the eagle, half eagle, and quarter eagle, of $10, $5, and $2.50, respectively. There was free coinage of both metals, and the coins, whether of gold or of silver, were unlimited legal tender. There was thus established a true bimetallic system.

The original coinage ratio. The coinage ratio was fifteen to one.[1] This ratio was chosen, following a careful investigation and report by Alexander Hamilton, with the intention of making it as nearly as possible the market ratio. As it turned out however the market ratio was actually about 15½ to 1. In accordance with the principles with which the reader is now familiar, no gold was brought to the mint for coinage. Some silver was coined, but it happened that the new American dollars, though slightly lighter than the

---

[1] The gold dollar contained 24.75 grains of pure gold and 2.25 grains of alloy, a total weight of 27 grains of the standard metal. The silver dollar contained 371.25 grains of pure silver and 44.75 grains of alloy; *i.e.*, 416 grains standard. The coinage ratio was therefore $\frac{371.25}{24.75}$ or exactly $\frac{15}{1}$.

Spanish pesos, were found to be more attractive to the people of the West Indies and Central and South America because of their new and bright appearance. They therefore left the country almost as fast as they were coined, till their coinage was for this reason stopped in 1804. For a generation American coins made up only a small part of the monetary system. Foreign coins continued to circulate, though the main reliance was always upon paper money; *i.e.*, government notes and bank notes.

The coinage ratio changed. This state of affairs was not considered satisfactory, particularly the absence of American gold coins. Recognizing the cause in the discrepancy between the mint ratio and the market ratio, Congress in 1834 reduced the weight of the gold coins and so changed the legal ratio to about sixteen to one.[1] The market ratio remaining close to 15½ to 1, there was now a discrepency on the opposite side. There was no longer inducement to bring silver to the mint, the previous tendency to export silver dollars was strengthened, and it was profitable to hoard or melt down such few silver dollars as were in circulation. Gold on the other hand was brought to the mint, and American gold coins began to take their place in circulation. In 1837 certain slight changes were made in the coins, the purpose being to make all coins exactly nine tenths fine. The mint ratio, though slightly reduced, still remained practically sixteen to one (to be exact, 15.988 to one) and this ratio has continued to the present day.[2] Matters stood

---

[1] The exact details of the change may be shown thus:

	Gold Dollar			*Silver Dollar*			
Act of	*Pure Gold*	*Alloy*	*Standard Weight*	*Pure Silver*	*Alloy*	*Standard Weight*	*Legal Ratio*
1792	24.75 gr.	2.25 gr.	27. gr.	371.25 gr.	44.75 gr.	416. gr.	15 : 1
1834	23.2 gr.	2.6 gr.	25.8 gr.	371.25 gr.	44.75 gr.	416. gr.	16.002 : 1

There was no change made in the silver dollar. The legal ratio was now $\frac{371.25}{23.2}$ or $\frac{16.002\ +}{1}$.

[2] The exact results of the act of 1837 are shown in the following table:

	Gold Dollar			*Silver Dollar*			
Act of	*Pure Gold*	*Alloy*	*Standard Weight*	*Pure Silver*	*Alloy*	*Standard Weight*	*Legal Ratio*
1834	23.2 gr.	2.6 gr.	25.8 gr.	371.25 gr.	44.75 gr.	416. gr.	16.002 : 1
1837	23.22 gr.	2.58 gr.	25.8 gr.	371.25 gr.	41.25 gr.	412.5 gr.	15.988 : 1

thus for more than a generation, bimetallism being legally in force but the practical effect being, from 1834 to 1873, gold monometallism. After the middle of the century, following the gold discoveries in California and Australia, great quantities of gold were coined.

The end of bimetallism. In 1873 the standard silver dollar was dropped from the list of coins to be made at the mint. This action was taken as a mere detail of a general revision of the coinage laws. It attracted little attention at the time and had no practical effect, since there was still no desire to bring silver to the mint for coinage into standard dollars. Although there had been free coinage of silver from the beginning, there had actually been very little coinage of silver dollars, and practically none at all since 1834. As a matter of fact neither gold nor silver was then in circulation, both having been expelled, through the action of Gresham's law, by the still cheaper "greenbacks" of the Civil War, which did not come back to par till 1878.[1] No one had any interest in the obsolete silver dollar. The act of 1873 thus definitely put an end to bimetallism and made the monetary system in law what it had long been in practical effect, gold monometallism.

Change in the market ratio. Now, by one of those extraordinary coincidences which human affairs occasionally encounter, the act of 1873 had scarcely gone into effect when the market ratio between gold and silver, which for two centuries had remained almost constant (close to 15½ to 1), began a spectacular rise. The very next year it reached the legal ratio of 16 to 1 and passed it, four years later it was 18 to 1, and in 1894 it reached 32 to 1, just double the legal ratio. Various causes coöperated to bring about this remarkable result, among them the decline in the use of silver currency in India and a number of the European countries, the discovery of rich and extensive silver mines in America, and a rise in the value of gold itself.

If the act of 1873 had not demonetized silver (*i.e.*, abolished its free coinage), silver would unquestionably have been brought to the mint for coinage very soon after the market ratio had passed the 16 to 1 mark, in 1874 or 1875. Within a year or two a flood of silver dollars would have inundated the currency, gold coin would have disappeared, and practical silver monometallism would have prevailed. This statement of cause and effect is made with entire assurance, although it was vigorously denied by certain of the advocates of bimetallism, who claimed, right up to 1896 and 1900, that the corrective principle to which we have already given attention would have brought the market ratio back to 16 to 1 before all the gold had been expelled. As to this the lesson of American experience is conclusive. From 1792 to 1873 the average annual market ratio between silver and gold was never below 15 to 1 nor above 16¼ to 1. During this time the mint ratio was either 15 to 1 or (practically) 16 to 1. The average annual discrepancy was never as much as one point except in three years, and the maximum discrepancy was only 1¼, in the year 1813. Yet this small discrepency was too much to be corrected by the free coinage of the cheaper metal, and throughout the whole period there was in practical effect either silver monometallism or gold monometallism according as the legal ratio was below or above the market ratio. Certainly the restoration of the free coinage of silver could never have reunited the two ratios after their separation in 1874.

---

[1] See below in this chapter, under credit money.

European experience. Were it not for limitations upon our space, it would be interesting to examine at length the record of European experiences with the monetary standard. England is the traditional country of gold monometallism, her gold standard dating from 1816. In the rest of Europe the double standard generally prevailed for centuries, silver being the commoner metal. During the nineteenth century the maintenance of bimetallism involved various vicissitudes and was finally abandoned by all the leading nations within a year or two of the date which marked the demonetization of the silver dollar in the United States.

Bimetallism in France. Among all the countries of Europe, the experience of France is perhaps the most interesting and instructive. It will repay our brief attention. The French mint ratio was 15½ to 1, which was so close to the market ratio throughout the first three quarters of the nineteenth century that her bimetallic currency never entirely lost either metal until, at last foreseeing the expulsion of gold in 1873, she gave up the free coinage of silver. The French experience is interestingly recorded by Professor Taussig, as follows:[1]

"Whenever the price of silver fell in terms of gold, silver tended to be sent to France for coinage, and gold tended to flow out of France. Whenever the price of silver rose in terms of gold, gold tended to be sent to France for coinage, and silver tended to flow out. A high price of silver in terms of gold means, of course, a low market ratio, while a low price of silver means a high ratio. During the greater part of the period from 1820 to 1850, the price of silver was somewhat lower than the equivalent of the French ratio of 15½ to 1. Silver tended to flow into France; gold tended to flow out. The French circulation then consisted chiefly of silver; the proportion of gold was not large, and a very great substitution would have led to the complete disappearance of gold. That stage was nearly reached, but not quite. France was growing in population and wealth, and there was the basis for a large net increase in the stock of specie. Much of the added silver made its way into circulation without displacing gold, and the outflow of the latter metal, though it seems to have come very near to exhausting the stock in circulation, did not entirely do so.

"After 1850 the situation abruptly changed. The unexampled supplies of new gold from California and Australia were poured into the world's markets. The price of silver rose; the ratio fell. It became advantageous to send gold, not silver, for coinage into France. A very great influx of gold took place, amounting for the decade 1850–60 to over 3,000,000,000 francs ($600,000,000). A corresponding, though by no means an equal, outflow of silver took place. For in this period, as in that preceding, France increased her metallic stock, with the difference that now the addition was all in the form of gold, whereas before it had been chiefly in the form of silver. The silver which was steadily exported from France tended to keep down the price of silver bullion in the market, and so maintained the market ratio not far from 15½ to 1, though now with a tendency to a figure lower than 15½ rather than higher. . . .

---

[1] *Principles of Economics*, Third edition revised, pp. 271–274.

"Later in the nineteenth century another change set in, not quite so abrupt as that after 1850, but no less unexpected. The production of gold had reached its maximum about 1860, and thereafter barely held its own. The inflowing new supplies were still very great as compared with any period before 1850; but they spread over a larger area, and they were met by an increasing volume of goods. . . . On the other hand, a change began in the production of silver. Great discoveries were made in the United States, the beginnings of an increase in the productiveness of silver mining as striking as that which had taken place in gold mining. The price of silver in the market fell slightly about 1865. Silver no longer flowed out of France, and some silver flowed in. The market price for a few years was equivalent almost exactly to the ratio of 15½ to 1. Then in 1873 it fell more sharply, became equivalent to a ratio of 16 to 1, and led to a new inversion of the movement; gold began to flow out of France in large quantities, and silver began to flow in.

"This inversion proved unwelcome. Gold had come to be regarded, reasonably or unreasonably, as the preferable metal. The practice of England, the leading industrial country, was the main cause of this preference. The German Empire, when reorganizing its currency system in 1871, adopted the gold standard once for all, influenced chiefly by the English example. The coinage of the United States had been, after 1850, practically on a gold basis. France, not wishing to lose her gold, in 1873 stopped the free coinage of silver."

Conclusions regarding bimetallism. We are now in a position to formulate our general conclusions respecting the double monetary standard, or bimetallism. Bimetallism involves (1) a monetary unit which is defined as a certain weight of gold or a certain weight of silver, (2) free coinage of both metals, and (3) the legal tender quality for the coins of either metal. When the market ratio between silver and gold is the same as the legal ratio, both metals will be coined and both gold and silver coins will be in circulation. A discrepancy between the two ratios tends to stop the coinage of the metal which is relatively undervalued at the mint and to cause the disappearance from circulation of the coins previously made of that metal. The increased use of the other metal tends however to raise its market value and, if the discrepancy between the two ratios is quite small, may thus bring back the market ratio to agreement with the mint ratio before all of the undervalued metal has disappeared from circulation. A new equilibrium is thus created, with relatively less of the undervalued metal and more of the overvalued metal in circulation than before. A greater discrepancy between the two ratios cannot thus be overcome, but will continue, leading to the complete expulsion of the undervalued metal and the sole use of the other as standard money. Exactly how great a discrepancy between the market and mint ratios can be overcome by the added demand for the overvalued metal depends among other things upon whether the burden is borne by a great nation or a small one, or by few or many nations acting together. An international system of bimetallism including the majority of the leading nations, with agreement upon a mint ratio not far removed from the then existing market ratio, would undoubtedly have a powerful stabilizing effect on the market ratio. The experience of the United States shows that a discrepancy of one point is too much to be overcome by any one large and powerful nation. As a matter of fact, no nation has succeeded in maintaining permanently a system of bimetallism, and it is certain that no nation in the world could have maintained bimetallism for more than a year or two after 1873 at a legal ratio of 16 to 1 or less.

The "limping standard." When, in the period 1871 to 1874, the leading nations of the world definitely abandoned bimetallism and embraced the single gold standard, the former standard silver coins were left in a peculiar position. Their free coinage privilege was gone, and they were therefore no longer true standard money. At the same time they were generally left unlimited legal tender. Within a few years the value of silver had so fallen with respect to gold (*i.e.*, the market ratio had so risen) that the silver in these coins was no longer worth as much as the face value of the coins.

The coins no longer had full intrinsic value, as is the case with real standard money. Had free coinage continued, they would, as we have learned, soon have increased greatly in number and driven gold out of circulation. But now their quantity was fixed at the amount existing when coinage was stopped, and the tendency to drive out gold, which had already set in, was checked. These hybrid coins, neither true standard money nor yet real token coins since they were not redeemable in standard money, continued thus to circulate side by side with the standard gold coins and gave rise to the cynical title, the "limping standard."

The American silver dollar after 1873. Whereas in Europe, as we have seen, silver coins were generally more numerous than gold at the time when bimetallism was given up, in America it will be recalled that this was not the case. For more than a generation we had had virtual gold monometallism, and even before 1834 the coinage of silver dollars had been small and many of the coins made had been exported. In the whole history of the country up to 1873 the coinage of silver dollars had amounted to only $8,000,000. Silver dollars were thus in 1873 an insignificant part of our circulating medium, and the problem of the limping standard need never have been acute in America, had it not been for subsequent legislation of a highly artificial sort.

Reference has been made to the strong political demand for the restoration of the free coinage of silver at the legal ratio of 16 to 1 which developed in conjunction with the remarkable fall in the price of silver beginning in 1873 and continuing to the end of the nineteenth century. The success of this movement would have meant silver monometallism with a cheap silver dollar worth by 1894 only half as much as the gold dollar. This, as will be more fully explained in a later chapter, would have greatly benefited all debtors and equally injured creditors. The movement did not succeed, though it came very close to it, but it did cause Congress to enact two measures of a compromise nature which had a momentous effect upon the United States monetary system.

Bland-Allison and Sherman acts. In 1878 the so-called Bland-Allison act compelled the Treasury to purchase each month

$2,000,000 worth of silver bullion and coin it into "standard" silver dollars. One year's coinage under this law was three times the total amount of silver dollars coined up to that time, and the law remained in effect for twelve years, leading to the coinage of 378 million silver dollars. Then in 1890 the Bland-Allison act was superseded by the Sherman act, which increased the Treasury's monthly purchases of silver to 4,500,000 ounces. This silver was paid for by the issue of legal tender treasury notes redeemable in either gold or silver coin at the option of the Treasury. The silver so bought was simply held in storage by the Treasury, being coined into dollars only as required to redeem the treasury notes. This act remained in effect only three years. In consequence of it 187 million silver dollars were eventually coined, the last coinage being in 1905.

The result. The Sherman act, though differing in detail from the Bland-Allison act, had nevertheless the same fundamental effect. Both acts inflated the monetary system with silver dollars or notes backed primarily by silver dollars, whose metallic content was worth less than — by 1894 only half — the standard gold dollar. Such a movement, continued long enough, would have had the same effect as the free coinage of silver; *i.e.*, the expulsion of gold and the substitution of the silver standard. This effect was already being felt. Gold was rapidly disappearing from circulation, accompanied by heavy exports to foreign countries. The silver dollars and the treasury notes of 1890, while never formally at a discount, were more and more being regarded with suspicion, and the Treasury was hard put to it to maintain the stability of the monetary system and meet its gold obligations. The monetary inflation and weakness constituted one of the chief causes of the disastrous panic of 1893. Inflation was finally stopped by the repeal of the Sherman act in 1893, but the harm had already been done, and in 1894 and 1895 the United States Treasury barely escaped suspension and was forced to resort to extraordinary means to maintain its disappearing gold reserve.

The free silver party was defeated in the presidential elections of 1896 and 1900, and in the latter year Congress definitely established the gold standard. The heritage of the nineteenth century was a total coinage of 578 million silver dollars, almost all under the artificial stimulus of legislation subsequent to the demonetization of silver. The silver dollar is still in an anomalous situation. It has one of the attributes of standard money, being unlimited legal tender. It lacks the other attribute, free coinage, and it has not full intrinsic value. Its value is maintained equal to gold by its limited quantity and by the legal obligation of the Treasury to maintain all kinds of money at a parity, which implies, though it does not in so many words require, redemption in gold. The stock of silver dollars at present (December 1, 1925) is 523,671,078, of which only 55 millions are in circulation while 388 millions are represented in circulation by silver certificates (to be described later). No useful purpose is served by the silver dollars that would not be equally accomplished if they were made token coins specifically redeemable in gold. In that case the fifty odd millions required for circulation would remain, the silver certificates would be redeemed in gold, and the silver dollars held back of them could be sold for bullion or used for the making of silver token coins.

A lost opportunity. At the time of the World War the United States lost an excellent opportunity to get rid of the " limping " silver dollars without loss. The sudden extraordinary rise of all prices finally brought the price of silver to the point where the market ratio was less than 16 to 1. For several months in the winter of 1919–1920 there was thus more than a dollar's worth of metal in a silver dollar, and there was profit in melting down the silver coins. Some were melted and exported. Moreover there came a special demand for silver from the British and Indian governments. Congress, through the Pittman act of 1918, permitted the Treasury to melt and sell the silver dollars held back of the silver certificates up to 350 millions, the corresponding silver certificates being replaced by a special issue of federal reserve bank notes. A little over 270 million silver dollars were so melted down. But the Pittman act provided that as soon as the emergency was over the silver dollars must be replaced by new dollars coined from American silver purchased by the government. This has now been done, and we are back where we were with the heritage of the Bland-Allison and Sherman acts.

Fiduciary money. Of the three groups into which we found it convenient to classify money, our attention has so far been directed to standard money. We may now turn to the second group, fiduciary money, and study in turn the three forms, (*a*) representative money, (*b*) token money, and (*c*) credit money.

Representative money. The subject of representative money is fairly simple, and a brief discussion will suffice. As already defined, representative money consists of certificates testifying to the fact that the government holds standard money which will be delivered to the bearers of the certificates on demand. Such certificates are printed by the government and given to the people or the banks in exchange for standard money. The government must always hold for their redemption, and available for no other use, a quantity of the standard money exactly equal to the total of the certificates outstanding. The certificates are thus economically and legally simple warehouse receipts, similar to the warehouse receipts given for cotton or wheat in storage. Representative money makes no net addition to the total quantity of money in circulation. Its sole purpose is to make possible the holding and circulation of the standard money by means of these evidences of property rights rather than by the handling of the actual coin, thereby adding greatly to the public convenience.

Since the holder of representative money can always obtain standard money on demand, and since the government will also give the certificates freely in exchange for standard coin, there can never be an appreciable difference in value between representative money and the standard money. Depreciation of the former could occur only in case of default or the fear of default on the part of the government.

Representative money is not normally legal tender. Although of equal value to the standard money it is after all not standard money. He who has sold goods or rendered services in exchange for money is presumed to be entitled to the best money there is. If the representative money is acceptable, as of course it generally is, nothing is to be gained by making it legal tender. If however the creditor should prefer to receive standard money, the debtor is the one to go to the slight trouble of exchanging his representative money for standard money.

United States gold certificates. As an example of representative money we may cite the gold certificates of the United States. The contract printed upon every such certificate reads: "This certifies that there have been deposited in the Treasury of the United States of America ten [or some other number] dollars in gold coin payable to the bearer on demand." This is evidently the statement of a warehouse receipt. Gold certificates were first authorized by Congress in 1863, and they have since then become an important part of the United States monetary system. Indeed so fully is their convenience appreciated that the quantity in use, outside the Treasury itself, is more than double the amount of gold coin so used. On December 1, 1925, there were 1,691 million dollars of gold certificates outstanding and only 731 millions of gold coin in use outside the Treasury. If from these figures we deduct the sums held by the federal reserve banks and agents, we find that the amounts in actual "circulation" were on that date 1,109 millions and 426 millions respectively, which represent per capita $9.67 of gold certificates and $3.71 of gold coin in circulation. The United States Treasury held on this date, as a trust fund for the redemption of gold certificates, the full amount of 1,691 million dollars in gold coin or bullion. In spite of the principle laid down in the previous paragraph, the gold certificates, though originally not legal tender, were made so by a law enacted in 1920.

United States silver certificates. The silver certificates of the United States have precisely the same relation to silver dollars as exists between the gold certificates and gold coin, and here again it is evident that the people prefer the certificates to the actual coins. While there were on December 1, 1925, 524 million silver dollars, only 69 millions were outside the Treasury and only 55 millions in actual circulation, whereas 444 millions were represented by certificates, of which 388 millions were in circulation. The per capita circulation of silver certificates was $3.38, of silver dollars only 48 cents. Owing to the fact that the United States silver dollars are not standard money according to our definition, the silver certificates do not quite fit the definition of representative money. This anomalous situation is due to the peculiar historical causes with which we have already become acquainted.

Token money. We have defined token money as coins containing less than their nominal value of metal and redeemable in standard money. These coins are ordinarily of smaller denominations than the monetary unit, and their purpose is to facilitate payments involving fractional amounts. They are usually made of silver or of other less precious metals, as copper, nickel, or bronze. When made of the same metal as the standard coins, the token coins are of proportionally less weight. Thus there is not as much silver in two half dollars, four quarters, or ten dimes, as in a silver dollar (the latter having formerly been a standard coin).

The value of token money. It should be evident that token money does not derive its value from the material of which it is made. There is not ten cents worth of silver in the dime, nor is the metal in the five cent piece worth five cents, nor the bronze in the one cent piece worth one cent. These coins maintain their face value because by law they may always be redeemed in standard money at the United States Treasury, the only requirement being that they be brought in multiples of twenty dollars. Conversely the Treasury will always give out the token coins on request in exchange for standard money. The value of the token coins is thus determined and can never be either more or less than the value of the standard coin.

Quantity automatically regulated. This arrangement also determines the quantity of the small coins, keeping it just sufficient to meet the needs of the people. When for example there is a shortage of half dollars in any community, the merchants and others, finding themselves embarrassed in making change, send to their banks for more of these coins. The banks, seeing their half dollars running low, send to the Treasury for more, and it is the business of the United States Mint to have always a sufficient stock to meet such demands. In the opposite case, when there is a surplus of any small coin, the merchants find their tills piling up with this coin. They deposit the surplus with their banks, and the latter if necessary present the small coins to the Treasury and demand standard money in exchange.

Why token coins are light weight. It may be asked, why all this machinery for the small change coins? Why not have the small coins standard money of full intrinsic value and avoid all the complications of token money? The first and obvious answer is that gold, the universal standard metal today, is so valuable that the microscopic small change coins made of it would be quite unusable. We do not have even a gold dollar, for this reason; imagine a gold cent! However the question goes deeper than this. Why in a bimetallic system should the small coins not be standard silver of full weight? Or, even with the gold standard, why not put enough silver or copper or other metal into the small coins to give them full intrinsic value? American monetary experience furnishes an interesting inductive answer.

The reader will recall that the first United States coinage system, adopted in 1792, was based on bimetallism at the ratio of 15 to 1. Besides the then gold coins (eagle, half eagle, and quarter eagle), there were provided five coins of silver: dollar, half dollar, quarter dollar, dime, and half dime. These were all standard coins, full weight, free coinage, and legal tender. There were also, it should be stated, a cent and a half cent of copper. Now it will also be recalled that the mint ratio of 15 to 1 undervalued gold, so that only silver was brought to the mint for coinage. No serious difficulty arose with respect to the small silver coins, till the mint ratio was changed in 1834 to about 16 to 1. This undervalued silver, and, as we have already noted, silver dollars were no longer coined and such as were in existence tended to be melted down or exported. Exactly the same results began to occur in the case of the small silver coins, and in a short time the people were experiencing difficulty in making change. It took the country nearly twenty years to learn the lesson, but finally in 1853 Congress solved the problem by stopping the free coinage of fractional silver coins, reducing their weight to about ninety-three per cent, and providing that they should thereafter be coined from metal bought by the government and be issued to the people as needed. Thus was learned and applied the principle that in a bimetallic system the token coins of either standard metal must be light weight, lest at any time when that metal is undervalued by the mint ratio the small coins disappear from circulation. Under the gold standard it is likewise impossible to have small coins of other metals of full intrinsic value. The market value of copper, for example, varies greatly. If at any moment the cent contained exactly one cent's worth of metal, we may be sure that this equilibrium would soon be disturbed. Either copper would rise, in which case the coins would disappear from circulation, or copper would fall and the coins would no longer have full intrinsic value. Whatever the standard money therefore the token coins are properly of less than full intrinsic value and not open to free coinage; their value depends on their redeemability.

Limited legal tender. With respect to the other characteristic of standard money, legal tender, it has also been found convenient to make a departure in the case of token money. Standard money is unlimited legal tender. But it would be most awkward if one were compelled to accept in payment of a hundred dollar contract, four hundred quarters, or ten thousand one cent coins. On the other hand the debtor must be permitted to pay odd amounts in the fractional coin. Hence token money is made limited legal tender. For example the subsidiary silver coins of the United States are legal tender for amounts not exceeding ten dollars; the five cent and one cent coins, for amounts not exceeding twenty-five cents.

Summary. To sum up, a normal system of token money is marked by the following characteristics: (1) not open to free coinage, but coined on government account from metal bought by the government, (2) light weight; *i.e.*, less than full intrinsic value, (3) limited legal tender, (4) freely redeemable in standard money and freely obtainable for standard money.

Credit money. As already defined, credit money consists of notes by which the government or certain banks promise to pay standard or some other kind of money to the bearer. For example the United States notes, or "greenbacks," are the promissory notes of the United States Government. Upon each note is printed this agreement: "The United States of America will pay to the bearer five [or some other number] dollars." We have also familiar examples of bank notes. A national bank note is the promissory note of a national bank and contains this agreement: "The First National Bank of Chicago [or some other national bank] will pay to the bearer on demand ten [or some other number] dollars." These are simple promissory notes. Neither the government nor the issuing bank is required to hold standard money equal in amount to the credit money outstanding, which distinguishes credit money from representative money. Whether specifically so stated or not on the note, payment is normally made freely upon the demand of the note holder. For this purpose the issuing government or bank holds a "reserve" of standard or other specified kind of money (termed in America "lawful money"), so as to be always in a position to pay its notes as presented. The amount of this reserve in relation to the volume of notes outstanding may be specified by law or may be left to the administrative discretion of the government or bank of issue.

Under the normal conditions as thus stated, credit money will, like representative money, have exactly the same value as the standard money for which it may be exchanged. Also like representative money it will normally not be legal tender.

Irredeemable credit money. It must be confessed however that these normal characteristics of credit money have in actual history been honored as much in the breach as in the observance. Governments and banks have frequently "suspended payment" of their notes, and such suspension, particularly upon the part of governments, has often been long continued. Sometimes the bank or government may fail entirely and cease to exist, in which case payment may be made of only a part of the amount called for or may never be made at all. Bank failures, with whole or partial loss to note holders, are no rarity in monetary history, and we need not go outside America for an example, the Confederate States of

America, of a government which ceased to exist and whose promissory notes were never paid. The situation is quite distinct from that of representative money, where the government has always on hand the full amount of standard money not lawfully available for any other use, and where the danger of suspension is therefore negligible.

When credit money is not actually being paid on demand, the relation which normally maintains its value equal to that of standard money is broken. Even in such case a limited quantity of credit money may keep its full face value, particularly if there is hope of prompt resumption of payment. But this is the exception. As a general rule it is current redemption on demand which alone has power to keep the value of credit money at par with the standard money. Otherwise credit money will ordinarily depreciate, and the depreciation may be anywhere from a few per cent to loss of all value. Material depreciation of the credit money will tend, in accordance with Gresham's law, to drive the standard money out of circulation. The credit money then becomes the basis of the monetary system, its value no longer controlled by the value of the previous standard money. Indeed as we have defined the terms, credit money ceases to be credit money and becomes fiat money when redemption has definitely failed.

Fiat money. As previously defined, fiat money is money which has no intrinsic value itself and which is not redeemable in any other kind of money. In the case of all other species of money, value is imparted directly or indirectly by the value of some commodity. Standard money has full intrinsic value, and the several kinds of fiduciary money derive their value from the right they give the bearer to demand standard money either immediately or by a roundabout process, as in the case of the United States silver certificate. Fiat money bears no such relation to the value of any commodity. It may at first sight be hard to understand how such money can have any value at all, or why anyone should consent to receive it. Yet that this does occur is demonstrated over and over in the annals of the world's monetary history. Fiat money has almost always been degenerated credit money.

The people having become used to the credit money, force of habit continues its use even when redemption has ceased and the chance of future redemption has become virtually nil. Its declining value is evidenced by rising prices, as will be more fully studied in a later chapter, but it continues to circulate. Each individual feels that he can accept the money so long as he can pass it on to others, and so the one essential characteristic of money, general acceptability, is maintained. It is remarkable how far habit will thus lead a people. The continental bills of credit of the American Revolution circulated till they were worth less than one cent on the dollar. The common people of Germany after the World War were still using their paper mark when its ratio of depreciation was impossible to calculate and when it was finally stabilized at one trillion to one!

Experience with credit and fiat money. Credit money is a regular part of the monetary system of every important nation, and monetary history is replete with innumerable experiences with this kind of money. Fiat money also has flourished in various parts of the world and at various times. During the World War and for several years thereafter nearly every great nation of Europe was on a fiat money basis; France, Italy, and some others are still there. Interesting though this field is, the limits of this book will not permit us to enter it, except as regards our own monetary history, where however we shall find examples of every conceivable variety of experience with both credit and fiat money to illustrate the economic principles which have been laid down.

American money in the colonial and Revolutionary period. Long before the Revolution the American colonists had become familiar with credit money and had generally so saturated their currency system with notes of the several colonies and of certain so-called banks that credit money, or more accurately fiat money since redemption in standard money was the exception rather than the rule, had become their ordinary medium of exchange. The War of the Revolution was financed principally by the issue of the promissory notes of the Continental Congress, which soon flooded the circulation, driving out coin and becoming the regular money of the people. These notes purported to be credit money, the promise of the Continental Congress to pay coin. But there was no coin for that purpose, the notes were from the first not redeemable on demand, and no one knew when they would be or if they would ever be redeemed. They were thus from the first genuine fiat money. They were issued in large quantities; their depreciation was prompt and disastrous. Eventually, having declined to less than one cent on the dollar, the people lost all faith in them and refused to accept them. Gold and silver coins then gradually found their way back into circulation. In the financial reconstruction after the adoption of the Constitution the notes were redeemed, so far as presented, at the rate of one cent on the dollar. Thus disastrously ended our first national experience with fiat money.

From 1789 to the Civil War. The issue of credit money by the states, freely resorted to during the colonial and Revolutionary period, was forbidden by the Constitution of the United States. Notes of the national government continued to be issued at various times down to the Civil War. These issues, though not always redeemable at the outset and frequently depreciated, were all eventually redeemed at par, having never reached the low state of the Continental bills of credit. During this period the main reliance for currency was upon another form of credit money; namely, bank notes. A strong government bank was in existence from 1791 to 1811 and another from 1816 to 1836. The notes of these United States banks were real credit money, currently redeemable and of high standing. Local state banks also issued their notes, some redeemable on demand and as good as gold, others of varying degrees of irredeemability. During the period 1811 to 1816 and for a short time following 1836, when there was no United States bank, the country was on a fiat money basis, owing to the character of the local bank notes and the weakness of the government notes.

The greenbacks. The Civil War was the occasion of a large issue of notes of the United States, destined to become a permanent part of the country's monetary system. These United States notes, or " greenbacks " as they soon came to be called, were issued in order to meet in part the cost of the war. They were made full legal tender except for (1) payment of duties on imports (tariff duties) and (2) payment by the government of interest on the public debt. A statement to this effect may be found engraved upon any greenback to this day. The explanation of this peculiar legislation is interesting. Congress desired to force the new notes into general circulation and feared that, unless made legal tender, they might be refused by the people, though in this fear they were probably mistaken. However after long debate and much misgiving, the legal tender clause was adopted. But at this time the government was seeking to borrow for the needs of the war by the issue of its bonds, upon which it promised to pay interest in gold. It was clear enough that if the government might now claim the right to pay its bond interest in the new legal tender notes, its ability to borrow would be serious impaired. Hence the second exception to the legal tender clause. Congress had next to consider where the government could get the necessary gold, and so by the exception with respect to tariff duties it provided one source of gold income.

An example of fiat money. The greenbacks, though promissory notes of the United States, were not at the outset redeemable in standard money. It was generally supposed that the notes would be redeemed in gold coin as soon as the war was over, but this was not finally done until January 1, 1879. From the start the greenbacks depreciated and, through the operation of Gresham's law, drove most of the gold and silver coin out of circulation. The other kinds of money in circulation were either other classes of notes of the United States or the notes of banks, themselves redeemable in greenbacks. The greenbacks were thus the basis of the monetary system and, being irredeemable and depreciated, they gave the country a pure fiat money system for the seventeen years from 1862 to 1879, though always somewhat supported by the hope of ultimate redemption. So low did they fall (the lowest point was about thirty-five cents on the dollar) that the small silver coins, though reduced in weight seven per cent in 1853 as we have learned, became more valuable as bullion than as coin and generally disappeared from circulation. Congress found it necessary to provide paper token money (notes) which should not be more valuable than the greenbacks themselves and which soon acquired the euphonious title of "shin plasters."

Later history of the greenbacks. We shall not be able to go much further into the interesting history of the United States notes. On January 1, 1879, the Treasury began to redeem the notes in gold on demand, and redemption has continued ever since, albeit with considerable difficulty during the troublous times of 1893–1895, to which we have previously referred. At various times after the Civil War some of the notes were paid off, but in 1878, the year before resumption of specie payment, Congress forbade further retirement and required that as the notes were redeemed they should be paid out and so kept in circulation. This fixed permanently the amount of the greenbacks at $346,681,016. For the redemption of the notes the Treasury maintains a reserve of gold coin, which since 1900 has been fixed at $150,000,000. If ever redemption of the notes should bring this gold reserve below $100,000,000 the Treasury is compelled to bring it back to $150,000,000 by borrowing gold through an issue of bonds.

Credit money in the United States today. Credit money today forms a very important part of the monetary system of the United States. Besides the 347 millions of United States notes, there are (December 1, 1925) more than two billion dollars of the notes of the federal reserve banks and 711 millions of national bank notes. Of the total amount of money in actual circulation (*i.e.*, not held by the Treasury and federal reserve banks and agents) on December 1, 1925, amounting to 4,972 million dollars, 306 millions were in United States notes, 1,748 millions in notes of the federal reserve banks, and 667 millions in national bank notes. Altogether 2,721 million dollars, more than half of the total money in circulation, was credit money. The special problems connected with bank note currency will be more fully investigated in the following chapters on banking.

CHAPTER XXII

THE PRINCIPLES OF BANKING

Financing production. Let us for the moment assume a very simple type of productive organization, consisting only of manufacturers, wholesalers, retailers, and consumers. Goods pass in a direct line from the manufacturer to the wholesaler, to the retailer, and to the consumer. There are no banks or other financial houses to assist any one of them, and therefore all the financial burdens must fall on the manufacturers, the middlemen, or the consumers.

When the wholesaler buys goods from the manufacturer he pays cash for them; the retailer and the consumer do likewise. The financial functions of these seem slight, but are they? The manufacturer is compelled to use his own funds in paying for materials, labor, etc. while he is manufacturing the goods; the wholesaler ties up his resources in a stock of goods until he has disposed of them to the retailer, and the retailer is in the same position. Even the consumer may buy somewhat in excess of his immediate needs. As a matter of fact, each one is *financing* himself.

Some of the consumers, let us assume, find it impossible to pay for the goods immediately, and they induce the retailer to " trust " them for small amounts. But as a result of this situation the retailer may find himself unable to pay cash for the goods he buys from the wholesaler. He is confident that, if he were given three months in which to make payment, he could sell a certain amount of goods and secure payment from his customers, and so he suggests to the wholesaler that he be allowed to purchase the goods on the basis of deferred payments — that he be given three months' credit. If the wholesaler agrees to this, he assumes the burden of financing the retailer as well as himself. Obviously to do this he must have fairly large resources or be compelled to curtail the extent of his operations, provided he still pays the manufacturer cash on delivery of goods. But he may make the same plea to the manufacturer which the retailer made to him, urging in extenuation his treatment of the retailer. If the manufacturer accedes to his request, he is financing himself, the wholesaler, the retailer, and the consumer. He is selling goods and receiving in exchange a written or verbal promise to pay for them sometime in the future.

Credit. Here we see the essence of the credit system. Credit involves the purchase of something in the present — merchandise, money, or services — agreeing to pay for it in the future. In business the amount and time of the future payment is usually determined and the payment is almost always made in money. But this is not necessarily so. A farmer may borrow a few bushels of seed wheat promising to return a certain number of bushels of wheat when his crop is harvested. This is as truly a credit transaction as when a banker loans $1,000 to be repaid in money one month from the date of the loan, even though the repayment in the former case is to be in wheat and at a time which cannot be fixed exactly when the loan is made.

The credit system rests ultimately on a belief in the ability of the borrower or purchaser to make the future payment to which he has pledged himself, but its existence is conditioned on the presence of individuals or of a class in society who are willing to advance present goods, whether money, merchandise, or services, in return for a promise to pay in the future. It therefore depends on the existence of a surplus, over and above present needs, which can be placed at the disposal of others, and the extent of the credit system is determined by the volume of savings.

In our example credit has been entirely of the type we shall call *commercial credit;* that is, it has been extended by the manufacturer, the wholesaler, or the retail merchant, and this held true in the main of the industry and commerce of former times. The financial houses devoted exclusively or primarily to financing business men are of fairly recent origin. For although it is true that banks have existed for many centuries — the Bank of Venice was founded in the twelfth century — and that we find money lenders and financial houses of importance in the medieval period, they were the exceptions to the general rule, and the producers were forced to rely on their own or their families' resources or on each other. It was not until the eighteenth and nineteenth centuries that banks and other institutions so increased in number and resources as to become a vital part of the economic fabric, and that *bank credit* (a promise by a bank to pay) surpassed commercial credit in importance.

Introducing the commercial bank. During the history of banking there have developed several different kinds of banks, and at the present day we need to distinguish at least three types; namely, commercial banks, trust companies, and savings banks. Land banks might be added as a fourth class. It is however the commercial banks which play the predominant rôle in the modern economic organization, and it is to this class of banks that our attention will be chiefly directed. A few words at the close will suffice for the other types.

The reader may be presumed now to know that a commercial bank's business consists primarily in making loans to and receiving deposits from its customers; *i.e.*, in the two functions, or better the double function, of discount and deposit. Let us plunge at once into the middle of our subject by assuming a moderate-sized city bank and following through in detail from beginning to end some of its typical operations.

Example of the loan. Let us assume that M, a dry goods merchant in the town, is engaged in buying his stock of goods for the early spring trade. As is usual, he does not have enough capital of his own for the purchase of his entire stock but depends regularly upon the financial assistance of his bank. He therefore goes to the bank and asks for a loan, let us say, of $50,000. The officers of the bank, after satisfying themselves as to his character and the state of his business, agree to accept his promissory note for $50,000 payable, let us say, sixty days from date.

Discount. The next step is to calculate the interest on $50,000 for sixty days at the rate at which the bank is making loans on this day, say six per cent. This is $500. Deducting this sum, called the *discount*, from the *face* of the note leaves $49,500 as the amount which M is to receive, which is called the *proceeds*.[1] The term "discount" is used both to refer to this interest deduction and also to describe the whole process of making a bank loan. When used in connection with banking, it may be defined thus: *Discount is either* (1) *the deduction of interest made by a bank, or* (2) *the process of making a bank loan.* No ambiguity should arise from the two senses in which this word is used.

Proceeds. The next question is as to the form in which M shall receive the proceeds of the loan. There are ordinarily three forms, (1) money (other than the bank's own notes), (2) the notes of the bank, (3) deposits. The first is simple enough and needs no explanation. In the second case the bank gives M, in exchange for his promissory note, its own promissory notes. It is a swapping of notes. But the notes of the bank are money; they are generally acceptable throughout the community, and M has received what he wanted, the means for making immediate payment for his goods. Indeed the ordinary business man does not distinguish between different kinds of money; to him there is practically no distinction between the first two cases. To the bank, as we shall see later, there is a very real distinction.

Both of these however are the exceptional cases, at least in the United States. In the great majority of instances the borrower receives the proceeds of his note, not in money ("cash"), but in the form of a deposit credit, and this brings us to the interesting and important subject of deposits. Everyone has a clear enough idea of the process by which money may be taken to a bank and left — "deposited" — for safe keeping. The bank credits the customer's account on its books, and the latter obtains the right to withdraw the specified amount of money, or any part of it, at will. Now in the case of a loan the amount of the proceeds is (in the ordinary case) credited to the account of the borrower as a deposit, precisely as though he had brought money. If it is any aid to visualizing the process, it may be imagined that the customer received the proceeds in money and immediately stepped over to the receiving teller's window and deposited the money. Actually however no money has figured; this is indeed not a money transaction. It is once more a swapping of credit. The bank receives the customer's promissory note and gives him in exchange the right to receive money from it on demand, which is precisely what he would have if he had taken instead the bank's notes.

---

[1] It will be noted that this calculation is not precisely true discount. M has received from the bank $49,500 for 60 days, at the end of which time he must pay $50,000, being the $49,500 plus interest upon $50,000. The process of true discount (as was fully explained in Chapter XIX) would be to determine the present worth of $50,000 payable in 60 days, which is $49,504.95, the discount being $495.05. The simpler method followed generally by the banks is appropriately called "bank discount."

Unsecured note. We must now retrace our steps in order to consider the several forms which the ordinary commercial loan may take. There is first the *unsecured note;* that is, the simple promissory note of the borrower. If signed by the borrower alone, it is called "one-name paper." However, since banks are usually reluctant to loan upon an unsecured one-name note, it is customary for the borrower to obtain the signature of some other person in addition to his own. The signature of this other person makes him liable for the payment of the note in case the borrower, or "maker," should default. A note so signed is called "two-name paper." There may even be "three-name paper," containing two signatures besides that of the maker.

Secured note. We have next the *secured note.* This is likewise the promissory note of the borrower, but accompanied by "collateral security." This means that the borrower delivers to the bank along with the note certain property of his, worth at least as much as the amount of the note, with an agreement that in case the note is not paid when due the bank may sell the security and so reimburse itself for the note. Stocks and bonds of corporations and government bonds are the property most used as collateral security, though warehouse receipts for certain commodities may be used. The secured note is one of the most common forms of bank loan.

Endorsed note. In the example which we used, M's loan might have been obtained by means of either an unsecured note or a secured note, whereupon he obtained the $49,500 from the bank and then paid the wholesaler. The purchase of his stock in trade might however have been somewhat differently financed. It is common practice for the wholesalers to sell to their customers on credit; that is, to accept their promissory notes in place of cash. Suppose that M has such an arrangement with a wholesaler, N, in the same city. He will give N in exchange for the goods his promissory note for $50,000 payable in 60 days. The wholesaler, N, being himself actively engaged in business with all his capital, will not be willing to hold M's note till maturity; he needs the cash before that. N therefore avails himself of the services of the bank. He first *endorses* M's note; that is, signs his own name, thereby making himself liable for its payment in case M should not pay. He then takes the note to the bank, which discounts it precisely as in the case when M borrowed on his own note. The interest, or discount, is calculated and deducted from the face of the note, and the proceeds are given to N, either in money or, more likely, as a credit on his deposit account. The note becomes the property of the bank, and when it is due M will pay, not N, to whom he gave the note originally, but the bank. This gives us a third form of loan, the *endorsed note.*

Draft or acceptance. The bank loan may take still a fourth form; namely, the *draft* or *acceptance.* The retailer, instead of giving the wholesaler his promissory note in return for goods received, may give him the authority to *draw* upon him. This, in the case we are using for illustration, would mean that when N sent the goods to M he would write out an order upon M directing him to pay the $50,000 in sixty days to N's bank. This *draft* would then be taken to M, who would *accept* it; that is, write the word " accepted " and his signature across its face, which is his acknowledgment of his obligation to pay at the maturity of the draft. A draft when thus accepted is called an *acceptance.* It is two-name paper, since the drawer, N, is liable as well as M, upon whom it is drawn. N then takes the acceptance to the bank and discounts it precisely as has been explained in the case of M's promissory note discounted by N. At the maturity of the draft, M will pay the bank. These various forms of business men's notes and drafts are all called *commercial paper.*

A bank's balance sheet. These matters may be made clearer and the foundation laid for our further investigation, if we now give some attention to the balance sheet of the bank. A commercial bank is ordinarily a corporation, and to it may be applied what has already been learned about the organization and accounts of corporations. Avoiding details not essential to our present purpose, we may say that a bank is organized about as any other business corporation (see Chapter VI). The first balance sheet might read somewhat as follows:

STATEMENT OF THE CITY BANK, NOVEMBER 1, 1925

Assets		*Liabilities*	
Cash	$400,000	Capital stock	$400,000

The bank proceeds at once to buy land and a banking building, furniture and supplies, etc., for $60,000, and thereafter its balance sheet reads as follows:

STATEMENT OF THE CITY BANK, NOVEMBER 10, 1925

Assets		*Liabilities*	
Real estate	$ 45,000	Capital stock	$400,000
Furniture, etc.	15,000		
Cash	340,000		
	$400,000		$400,000

A loan for cash. The bank is thus ready for its first customer, whom we beg now to introduce in the person of M, whose desire for a $50,000 loan has already furnished us some food for thought. As has been already assumed, M offers his promissory note, either with another signature or secured by collateral. It is for $50,000 payable in sixty days, and the bank's rate of discount is six per cent. The discount is $500 and the proceeds, $49,500. M elects to receive " cash " other than the bank's own notes. The effect of this transaction upon the bank's condition is shown in the following statement:

STATEMENT OF THE CITY BANK, NOVEMBER 12, 1925

Assets		*Liabilities*	
Loans and discounts	$ 50,000	Capital stock	$400,000
Real estate	45,000	Undivided profits	500
Furniture, etc.	15,000		
Cash	290,500		
	$400,500		$400,500

It will be noted that the $50,000 promissory note acquired by the bank is listed among its assets opposite the heading "loans and discounts," that the cash has been reduced by the payment of $49,500 to M, and that the difference between these amounts, the discount, appears as undivided profits. Obviously the effect on the bank's statement will be the same whether M borrows by means of an unsecured note or a note secured by collateral. Even in case M finances his business by giving his promissory note to the wholesaler or permitting the wholesaler to draw upon him, the final result is the same, the only difference being that in these cases (that is, the endorsed note or the acceptance) M's obligation is discounted at the bank by N, the wholesaler, instead of by M himself. In any case the bank comes into possession of M's obligation, which M must pay to the bank at maturity.

A loan for bank notes. On the other hand the statement will differ according to the form in which the proceeds are given to the customer. The statement above illustrates the case where the proceeds are taken in the form of money other than the bank's own notes. If the bank had given its notes the statement would have read thus:

Statement of the City Bank, November 12, 1925

Assets		*Liabilities*	
Loans and discounts	$ 50,000	Capital stock	$400,000
Real estate	45,000	Undivided profits	500
Furniture, etc.	15,000	Notes	49,500
Cash	340,000		
	$450,000		$450,000

A loan for deposits. If finally the proceeds have been credited to the deposit account of the borrower, the statement would have been as follows:

Statement of the City Bank, November 12, 1925

Assets		*Liabilities*	
Loans and discounts	$ 50,000	Capital stock	$400,000
Real estate	45,000	Undivided profits	500
Furniture, etc.	15,000	Deposits	49,500
Cash	340,000		
	$450,000		$450,000

3400
4950
450

3400
495
05
9

A normal bank statement. For the sake of our further illustrative bank statements, let us assume that during the next two months the bank had done a flourishing business, making many loans, correspondingly increasing its deposits and notes and profits, that a part of the latter has been transferred to surplus, and that the bank has made some investments in stocks and bonds. The statement, corresponding more closely to a normal banking business, may be assumed to be as follows:

STATEMENT OF THE CITY BANK, JANUARY 9, 1926

Assets		*Liabilities*	
Loans and discounts	$1,250,000	Capital stock	$ 400,000
Stocks and bonds	450,000	Surplus	75,000
Real estate	45,000	Undivided profits	12,327
Furniture, etc.	15,000	Notes	50,000
Reserve[1]	182,460	Deposits	1,405,133
	$1,942,460		$1,942,460

Bank deposits. Our attention has thus far been directed principally to the function of discount. We must now inquire somewhat further into the subject of deposit. *A bank deposit is a right to receive money from a bank, evidenced by an entry on the bank's books.* Deposits are generally, though not always, payable on demand. They are ordinarily the result of discount, as has already been explained and illustrated. Deposits may also be obtained by paying money to the bank, or by "depositing" a check drawn by another depositor (as will be made clear later), or in certain other ways which need not concern us here.

To illustrate the deposit of money, let us now suppose that a customer brings in $10,000 of money for deposit. The bank's statement immediately thereafter will read thus:

STATEMENT OF THE CITY BANK, JANUARY 9, 1926

Assets		*Liabilities*	
Loans and discounts	$1,250,000	Capital stock	$ 400,000
Stocks and bonds	450,000	Surplus	75,000
Real estate	45,000	Undivided profits	12,327
Furniture, etc.	15,000	Notes	50,000
Reserve	192,460	Deposits	1,415,133
	$1,952,460		$1,952,460

---

[1] The term "reserve" is ordinarily used to include the bank's cash on hand.

It will be noted that nothing on the bank's statement distinguishes between deposits obtained by paying money to the bank and deposits obtained by the process of discount. There is no difference. What the depositor has is exactly the same in either case, not money, but the right to receive money. Deposits are not, as is sometimes popularly supposed, " money in the bank." Banks never have on hand a stock of money equal to their deposits. That could hardly be expected, since, as we have seen, most deposits originate without any money being given in to the bank. Even when money is deposited, the money ceases from that moment to be the property of the depositor. It belongs to the bank and may be spent by the bank as it sees fit. The depositor has simply the right to receive a certain amount of money whenever he chooses to demand it of the bank.

Discount and deposit. In fact both discount and deposit, contrary to the meanings popularly attached to the terms " loan " and " deposit," are actually exchanges or sales, transfers of certain rights for other rights or for money. When M borrowed on his promissory note, he sold the note to the bank, among whose assets it appeared thereafter. In exchange he received a deposit, another kind of property right, which became his asset. When N, having accepted M's promissory note in payment for goods, discounted the note, he sold it to the bank in exchange for a deposit, or perhaps for bank notes or some other form of money. Money deposited at the bank is, so to speak, " spent " — paid for a deposit.

The double function of discount and deposit is ordinarily a swapping of credit — the person's promissory note or draft or acceptance for the bank's credit in the form of notes or deposits. If it be asked wherein lies the advantage of merely exchanging one form of credit for another, the answer is that what the bank's customer wants is a medium for making payments. His own promissory note will not be generally accepted; the bank's note will. He has exchanged a form of credit generally unknown and unacceptable for the well-known and acceptable credit of the bank. The same is true of deposits, since practically all the customer's payments may readily be made by check. Herein is the fundamental characteristic of bank credit and of the service rendered by the commercial bank.

The check. The right to receive money which a bank deposit gives the depositor is exercised by means of the check. If M desires to withdraw fifty dollars from the bank in order to make a certain payment or for pocket money, he makes out an order upon the bank, *i.e.*, a check, in the following form:

No. *731*

New Haven, Conn., *Jan. 9*, 1926

THE CITY BANK

Pay to the order of..............*Myself*......................

..............*Fifty*.........................Dollars. *$50.00*

J. C. MILLER

He presents this order to the bank, which thereupon gives him fifty dollars in money and deducts fifty dollars from the amount due him on its books.

If Mr. Miller (whom we have previously called M) wished to pay a doctor's bill of fifty dollars, he might obtain from the bank fifty dollars by means of a check like the above and then take the money to the physician. But a simpler method would be to order the bank to make the payment direct to the physician. For this purpose a check would be used in this form:

No. *732*

New Haven, Conn., *Jan. 9*, 1926

THE CITY BANK

Pay to the order of............*Robert Brown*..................

..............*Fifty*.........................Dollars. *$50.00*

J. C. MILLER

From these examples the nature of the check is apparent. *A check is a written order by a depositor directing his bank to pay money to some person.* The person to whom payment is to be made may be the depositor himself or the bank or any specified third party or the bearer. Checks are usually made out on printed forms like those in the illustration above, furnished by the bank, but such a printed form is not necessary.

Payment by check. Let us now follow the later history of the above check. The check is handed or sent by mail to Dr. Brown, who takes it to the bank and demands the money. Certain essential facts must be established before the paying teller of the bank will hand over the money. First he must be satisfied as to the genuineness of the check, which is usually sufficiently attested by the signature of the maker, with which the teller is familiar. Next it must be certain that Mr. Miller actually has as much as fifty dollars to his credit on the bank's books. The teller must also be certain that it is Dr. Brown himself who is asking for payment; if he is not personally known he will have to be identified. Before payment is made the check must be endorsed; that is, the payee, Robert Brown in this case, must sign his name on the back of the check. Endorsement is the payee's acknowledgement that he has received the money; it makes the check a valid receipt for the bank and for the maker. Having made the payment, the bank debits the amount to Mr. Miller's account, and eventually, when the account is balanced, the cancelled check will be returned to him to be kept as a receipt for the payment. The effect of this operation will be to alter two items on the bank's balance sheet, making it read as compared with the last statement (on page 420) as follows:

STATEMENT OF THE CITY BANK, JANUARY 10, 1926

Assets		*Liabilities*	
Loans and discounts	$1,250,000	Capital stock	$ 400,000
Stocks and bonds	450,000	Surplus	75,000
Real estate	45,000	Undivided profits	12,327
Furniture, etc.	15,000	Notes	50,000
Reserve	192,410	Deposits	1,415,083
	$1,952,410		$1,952,410

Instead of demanding the money, the payee may "deposit" the check. He endorses the check and takes it or sends it by mail or messenger to the bank, with the request that the amount stated be credited to his deposit account. The bank must be certain as to all the facts except identification, just as when money is demanded, and must be satisfied as to the genuineness of the endorsement. It will then credit the account of the payee (*e.g.*, Robert Brown) and debit the account of the maker (*e.g.*, J. C. Miller). The result of this transaction is exactly as though the payee had collected the money from the bank and then immediately deposited it. The balance sheet of the bank is not changed at all; no cash has been taken out, and the total amount due depositors is not affected by a transfer from one depositor's account to another's.

The check thus furnishes a means of making payments so convenient, safe, and generally advantageous that, in the Anglo-Saxon countries at least, it is availed of almost universally, not only by business men but by others. Nearly everyone above the class of the very poor and ignorant keeps an account at the bank. Money when received is regularly deposited at the bank, and the same is true of the far greater volume of receipts which come in the form of checks. Conversely fully four fifths of all payments are made by means of checks rather than with money. The bank thus becomes an agent for making payments between the people by the transfer of deposits from one to another.

Collection between banks. The matter has thus far been considered upon the assumption of one bank with which all persons concerned had their dealings. As a matter of fact in every place of any size there will generally be found two or more banks. Moreover the useful device of payment by check has not been permitted to remain a merely local convenience but has absorbed the great mass of payments between persons in different cities, states, and even nations. This introduces certain complications, which must now be investigated.

When one receives a check, the obvious thing is to take it to one's own bank and either deposit it or ask for the money. This is done regardless of whether the check happens to be drawn on this same bank or upon another. In the first case the matter is simply arranged, as has been explained. Even in the second case, the bank, though under no obligation to cash or credit the check, will ordinarily do so as an accommodation to its customer. Having given its customer the cash or credited his account, the bank thereupon becomes possessed of the check with its claim upon the other bank. Within twenty-four hours ordinarily the bank will present the check to the other bank and collect. This other bank of course then debits the account of its depositor who drew the check.

The first bank, in performing this service, accepts at its face value a check drawn by a stranger, the genuineness of whose signature it has no means of testing and the state of whose account in the other bank is wholly unknown to it. For protection it relies simply upon the endorsement of the check by the customer who presented it. His endorsement is his guarantee that the check is good and binds him to reimburse his bank if the check should prove not good. Checks sometimes pass through several hands before finally reaching the bank upon which they are drawn. Each person (individual or bank) who thus passes on the check is ordinarily required to endorse it, and each endorser makes himself responsible for the soundness of the check.

Thus the check system for payment of money or transfer of deposits operates when there are two or more banks in the town. Daily each bank sends a messenger to each of the others and collects for all checks of each other bank which it has paid. But in a large city with many banks this business of collecting by messenger becomes troublesome and expensive. In a city with fifteen banks there might have to be 15 × 14, or 210, separate calls each day. Sooner or later the point is reached where it is worth while to avoid this great labor and expense by organizing a clearing house.

The clearing house. The clearing house is an association of the banks of a city through which all the accounts between the banks for paying each other's checks and drafts are settled. The business of clearing takes place daily at a certain hour. Before that time there occurs in each bank the same preparatory operation. The clerks gather together all checks of the other banks which were paid the previous day and arrange them in bundles, one for each bank. A list is made out showing the amount thus due from each of the other banks with the total due from all the banks. Sometimes there is also made out a ticket for each of the other banks showing the amount due from that particular bank. Then two clerks, gathering together all the bundles, the tickets, and the list, proceed to the headquarters of the clearing house. Here there is a room containing a raised desk at one end for the governor and his assistants and on the floor a row or rows of desks, one for each member bank. Of the two clerks one takes his place behind his own bank's desk while the other, carrying the bundles of checks and the tickets, stands in front of the desk. The list showing the amount due this bank from all the others has been sent up to the governor's desk. In like manner, each of the other banks is represented by two clerks, one behind and one in front of its own desk.

At the exact moment for beginning the operation of clearing a signal is sounded on the governor's desk. The clerks in front then walk along the row of desks depositing at each the bundle of checks and the ticket against that particular bank. In a minute or two the checks and tickets have thus all been distributed, and it takes only a few minutes more for each desk clerk to add up the tickets or the sums on the outside of the envelopes and so ascertain the amount his bank owes to all of the other banks. These results are reported to the governor, whose assistants are able quickly to discover whether any error has been made. As soon as the figures, corrected if necessary, are compiled, the governor announces the result. Each bank now knows exactly how much it owes all the other banks and how much all the other banks owe it. These two sums are regarded respectively as a debt to the clearing house and a credit due from the clearing house. The balance represents an amount which the bank must pay to the clearing house or receive from the clearing house as the case may be. Thus by one payment either to or from the clearing house, each bank settles its accounts with all the other banks. Of course the amount payable by all the debtor banks to the clearing house is exactly equal to the amount payable by the clearing house to all the creditor banks, and the clearing house itself comes out even each day.

Clearing for outside banks. All the banks of a city do not necessarily belong to the clearing house. An outside bank will either make its collections by messenger and make its payments to messengers sent by the other banks in the primitive way, or it may arrange to clear through another bank which belongs to the clearing house. In this case the agent bank presents as its own the claims of the outside bank and accepts as obligations against itself claims of the other clearing house banks against the outside bank.

The volume of clearings. The volume of the business thus simply and smoothly effected by a large city clearing house is enormous. In the New York City Clearing House there were exchanged, during the week ending June 6, 1925, checks and similar obligations to the amount of $6,527,873,276. For the year 1924, the New York City clearings fell just short of 250 billion dollars. The New York City Clearing House stands head and shoulders above any other. For the whole of the United States the clearings in 189 cities, during the week ending June 6, 1925, were a little over 11 billions, and for the whole year 1924, over 455 billions. One can imagine what a gigantic task this would impose upon the banks were they still compelled to rely upon the primitive method of exchanging their obligations.

Out of town checks. Checks are regularly used for payments in other towns. When goods are purchased by mail order, payment is ordinarily made by a check drawn by the purchaser upon his own bank and sent by mail to the seller, it may be a thousand miles away. The maker of the check gives himself no further concern over the matter. The receiver of the check deposits it in his own bank, receives credit for it, and concerns himself no further. His bank has the task of collecting from the other distant bank. Thousands of payments are thus made every day, and every day the mails carry thousands of checks back and forth all over the country. There is thus imposed upon the banks a service of great importance and enormous volume. An elaborate system has been developed whereby this business is handled with swiftness and certainty and the minimum of expense and, by offsetting checks against each other, with remarkably little actual transfer of money.

Paying a loan. We must now return to the further study of the business of discount and deposit. We have used for illustration a loan of $50,000 for sixty days made at the City Bank by J. C. Miller. On January 11, 1926, the loan matures and must be paid at the bank. This is equally the case whether Mr. Miller borrowed directly from the bank on his own promissory note or gave his promissory note to the wholesaler, N, or accepted a draft drawn upon him by N, so long as the commercial paper has been discounted at the bank.

Mr. Miller now presents himself and offers to pay his note. He may pay in either of four ways, (1) in money, other than the notes of this bank, (2) in notes of this bank, (3) by check upon his deposit account in this bank, or (4) by a check upon an account in some other bank. The effect upon the bank's condition is readily determined. In any case the loans and discounts will be reduced by $50,000, representing the note paid and returned to Mr. Miller. If the note is paid in money the bank's cash, or reserve, will rise by $50,000, and the same result will follow, within a few days at most, if payment is made by a check on another bank, since this will be promptly presented to the other bank, through the clearing house or by mail or messenger, and collected. If the note is paid by giving notes of this bank, the note item on the liability side of the account will be reduced by $50,000. Finally if payment is by a check on Mr. Miller's account in this bank, the bank's deposits will be correspondingly reduced. Two changes of $50,000 each counteract each other, and the statement remains in balance. The reader can easily draw up a new statement representing the bank's condition after payment of the note according to any one of the possible assumptions.

The normal business of banking. Acquaintance has now been made with the more important operations which make up the normal daily business of the ordinary commercial bank. The bank's " portfolio " of loans and discounts contains hundreds of thousands of notes and acceptances of individuals and corporations, to which new notes are being added daily and from which notes are being as constantly withdrawn when they mature and are paid. The bank's deposits are being constantly added to as new loans are made or as customers bring in for deposit either money or checks upon other banks, while on the other hand withdrawals by check are continually tending to reduce deposits, particularly in connection with the payment of loans. The bank's notes outstanding tend likewise to rise and fall in harmony with the loans and discounts. From time to time securities held by the bank are sold or new ones purchased, with corresponding changes in the " stocks and bonds " item. In connection with many of these operations changes are continually occurring in the bank's reserve, as cash is either received or paid out. The business is normally profitable, and the " undivided profits " item gradually increases, reduced at intervals by the declaration of dividends or by transfer to surplus.

Function of the reserve. The deposits and notes of a bank represent claims against the bank held by outsiders and payable in cash on demand. The bank must be ready to pay any note holder or depositor any amount to which he is entitled at any time; refusal to pay means " suspension " or failure. For this purpose the bank must have always on hand a sufficient stock of cash, appropriately called the *reserve*. No commercial bank however ever holds a reserve equal to the sum of its demand liabilities. In fact the reserve will seldom be as much as twenty-five per cent and may be very much less. In the last illustrative statement (page 423) the bank's liabilities payable on demand (notes and deposits) appear as $1,465,083, for which there is a reserve of $192,410, a " reserve ratio " of a little more than thirteen per cent. Evidently a simultaneous demand from all depositors and note holders would wreck the bank. But of course no such simultaneous desire for payment is to be expected, and the bankers learn by experience about what demands may be anticipated from day to day. They make provision for a corresponding reserve with a proper margin for emergencies, and it is thus that the reserve ratio is determined except when legal requirements interfere.

Profits vs. safety. The banker is impelled by two counteracting motives, profits and safety. The bank derives its profits principally from the making of loans, and the larger its portfolio of loans and discounts the larger in general will its profits be. But as loans are made, deposits are created or the bank's note issues are increased. Deposits and notes on the one side and loans and discounts on the other side tend thus to rise and fall together. And as the sum of deposits and notes increases without corresponding increase in the reserve, the reserve ratio falls, and the bank's condition becomes proportionately less safe. To the banker's desire for profits is thus opposed the necessity of keeping a safe ratio between the reserve and the demand liabilities. The problem is solved differently according to the circumstances and the policy of the particular bank. If the bank's clientele consists mainly of conservative business men engaged in sound and safe enterprises, a relatively low ratio may be ample. If the bank is loaning for more venturous enterprises, it must safeguard itself by a larger reserve ratio. Again some bankers are by nature venturesome, "progressive," or reckless; they expand their business at the risk of safety. Others are cautious and "conservative" and keep their reserve ratios well above the danger line.

How the reserve ratio is controlled. There are three principal means of controlling the reserve ratio, (1) by the purchase and sale of securities, (2) by rediscounting loans, and (3) by changing the rate of discount.

Buying and selling securities. As we have seen, a bank usually invests a part of its resources in corporate securities, the main purpose being to derive some income upon funds which if kept as idle cash in the reserve would be earning nothing, while at the same time having something which may be converted into cash whenever needed. The bank therefore chooses for investment securities with a ready market which may be sold at short notice. Consider now the statement on page 423. The bank has a reserve ratio of a little more than thirteen per cent. If this is considered too small some securities may be sold for cash. Let us assume that securities carried on the books at $100,000 are sold for that amount, reducing the stocks and bonds to $350,000 and increasing the reserve to $292,410. The reserve ratio is now $\frac{292{,}410}{1{,}465{,}083}$, or almost exactly twenty per cent. The bank has sacrificed the future interest or dividends on $100,000 of securities, but it has made its position materially safer. Conversely a purchase of securities would reduce the reserve and consequently the reserve ratio. It is because of their availability thus to strengthen the reserve that a bank's readily convertible securities are sometimes called its "secondary reserve," though this figurative use of the term must not lead one to think that securities can ever be real reserve.

Rediscounting. The banker who wishes to strengthen his reserve ratio may instead of selling securities sell some of his loans. This is called *rediscount* and is precisely the same process as that by which the bank originally acquired the notes. Some investor is found, either another bank or a broker, willing to take certain promissory notes or drafts at a price determined by discounting them at a certain rate for the time that must still elapse before their maturity. Suppose our bank selects from its portfolio a note for $75,000, originally acquired two months ago and payable three months from its date, having therefore still one month to run. If another bank is ready to rediscount it at eight per cent, the City Bank will receive for it $74,500; that is, the face value less one month's interest at eight per cent. The result of this operation, starting from the statement on page 423, will be reflected in a new statement, as follows:

STATEMENT OF THE CITY BANK, JANUARY 11, 1926

Assets		*Liabilities*	
Loans and discounts	$1,175,000	Capital stock	$ 400,000
Stocks and bonds	450,000	Surplus	75,000
Real estate	45,000	Undivided profits	11,827
Furniture, etc.	15,000	Notes	50,000
Reserve	266,910	Deposits	1,415,083
	$1,951,910		$1,951,910

It will be noted that loans and discounts have been decreased by $75,000, the reserve is increased by $74,500, while the discount is reflected as a loss, decreasing the undivided profits by $500. The reserve ratio is now $\frac{266,910}{1,465,083}$, or something more than eighteen per cent, instead of thirteen per cent as it was before, showing that the desired result has been accomplished.

Control by the rate of discount. By purchase or sale of securities and by rediscounting, the banker is able to make a quick change in his reserve ratio. Of even greater significance however is the constant regulation by means of the rate of discount. As has been observed, in the course of a bank's daily business new loans are constantly being made and notes held are constantly maturing and being paid. There is thus an inflow and an outflow of commercial paper, which under ordinary circumstances may about offset each other, keeping the amount of loans and discounts fairly constant though the individual notes held are always changing. Now the payment of loans on hand is already determined by their respective maturities, but anything that could be done to speed up or slow down the rate of making new loans would have power sooner or later to change the amount of loans on hand. For accomplishing this purpose the banker has always an instrument at hand; namely, the discount rate.

Reverting once more to the statement on page 423, let us suppose that the banker decides that his reserve ratio, thirteen per cent, is too small. He has been making loans, let us say, at six per cent, and for some time the new loans have about offset those that were maturing. He therefore raises his discount rate to seven per cent. Since he is now charging a higher price for the service of making loans, fewer people will avail themselves of his services in accordance with the law of demand, and the making of new loans will fall off. The old loans continuing to mature as before, the net result will be a decline in the bank's loans and discounts. Furthermore note that when customers pay off their loans they must pay either (1) in cash or in checks on other banks which will promptly be converted into cash, or (2) in the notes of this bank, or (3) in checks on their deposit accounts in this bank. Since loans are being paid faster than new ones are being made, the bank thus finds its reserve increasing and its demand liabilities decreasing. The reserve ratio thus rises. Let us suppose that after a week's time loans and discounts have fallen off by \$100,000, cash has increased \$15,000, deposits declined \$80,000, and notes declined \$5,000. The result upon the bank statement, as compared with the condition on January 10, 1926 (see page 423), is as follows:

Statement of the City Bank, January 17, 1926

Assets		*Liabilities*	
Loans and discounts	\$1,150,000	Capital stock	\$ 400,000
Stocks and bonds	450,000	Surplus	75,000
Real estate	45,000	Undivided profits	12,327
Furniture	15,000	Notes	45,000
Reserve	207,410	Deposits	1,335,083
	\$1,867,410		\$1,867,410

The reserve ratio has been materially increased, being now $\frac{207,410}{1,380,083}$ or fifteen per cent, as against thirteen per cent. A continuation of the high discount rate will further tend to strengthen the bank's position.

Conversely a reduction of the bank's discount rate will stimulate business and cause new loans to come in relatively faster than old loans are being paid off. Since the proceeds of the loans must be given in cash or in notes or in deposits, the tendency will be toward a decline in the reserve and an increase in the demand liabilities, thus lowering the reserve ratio. The rate of discount is thus, as it were, a control lever, by which the banker regulates his business, steering a middle course between the desire for large profits and the danger of failure from inability to meet the demands of depositors and note holders.

Note issue. The reader who has any practical knowledge of banking will not have failed to observe that in this discussion we are concerned primarily with the development of the fundamental principles of banking rather than with the description of any particular banking system. This distinction appears with particular sharpness in connection with the subject of note issue. This banking function has been so much regulated by legislation in the different countries, particularly in the United States, that the reader may have been surprised at the somewhat casual way in which we have thus far treated bank notes as though they were about the same thing as deposits. As a matter of fact the essential identity of bank notes and deposits is in the fundamental nature of banking; it is only legislation that has sharply distinguished them. This point needs some further elaboration.

Notes and deposits: fundamental similarity. We have already defined a bank deposit as a right to receive money from a bank, evidenced by an entry on the bank's books. *A bank note is a bank's promise to pay money on demand.* The term "money" in both these definitions is of course exclusive of the bank's own notes. It will be noted that these definitions are almost identical. Deposits and notes are both forms of bank credit; both are liabilities of the bank; both are (except in the special case of time deposits and very rarely in the case of notes) payable in money on demand. When a bank makes a loan, it is immaterial to it whether the credit given the borrower is in the form of notes or deposits. To the borrower it is merely a question whether he shall receive the evidence of his right to receive money from the bank as an entry on the bank's ledger and in his own pass book or inscribed on little pieces of paper which he will take away with him. These differences are superficial, not fundamental.

Differences between notes and deposits. There are however real differences between notes and deposits. The right to receive money from a bank, if in the form of a deposit, may be transferred only by check. Now a check is of itself no obligation of the bank. The person who receives it must be sure (1) that the signature upon the check is genuine and (2) that the maker actually has at least that amount on deposit and that it will still be on deposit when he presents the check to the bank for payment. Otherwise the check is not good. Checks are thus available for use only among people who are acquainted with one another or at least have confidence in one another. Their availability is also limited to those who are in the habit of doing business with banks. Checks are thus not generally convenient for payments between strangers, for paying travelling expenses, or for making ordinary wage payments. Where they are available, checks are usually a safer and more convenient means of payment than money. But after all they are not money, since they are not generally acceptable.

The bank note on the other hand is money. It is the definite obligation of a bank and, barring the risk of counterfeit or of the failure of the bank, may be accepted without any question as to the honesty of the one who tenders it and may be held indefinitely. The bank note is generally acceptable; it may be used for all kinds of payments. Bank notes again are expressed in round sums, convenient for payments of all amounts. Although representing rights to receive money from the bank, bank notes may pass about for weeks or months or years before being presented to the bank for payment. They may travel far away from the bank of issue. The check on the contrary usually appears for the purpose of a particular payment, it is expressed in an odd amount of dollars and cents, and its life is ordinarily limited to that one transaction, after which it returns to the bank and is cancelled. That is, deposits require generally a new check for each separate transfer. These distinctions between notes and deposits, though not so fundamental as their similarities, are yet real and important and justify the issue of bank credit in the two forms.

An important corollary of the differences between notes and deposits is that the opportunity for abuse and fraud on the part of an unsound or dishonest bank is far greater in the case of notes than of deposits. The deposits of a commercial bank are extremely active. Depositors are frequently demanding money, and deposits are being continually transferred by check, each check being promptly presented to the bank. The bank's condition is thus subject to constant test by persons having more or less direct contact with the bank and knowledge of its management and condition. The reverse is true in the case of notes, and an unsound or dishonest bank may long continue to defraud the public by means of its notes.

Legal restrictions. This is in part the explanation of the prevailing attitude of government toward the banking functions of deposit and note issue. In spite of their fundamental similarity and their original status as common law rights, the legislation of most countries has drawn a sharp distinction between them. The banks have generally been permitted to handle their deposits with little interference by the government. The function of note issue has almost everywhere been subject to strict regulation. In most countries the right of note issue is restricted to certain kinds of banks or even to one particular bank. Thus in the United States only the national banks and the federal reserve banks issue notes. In Canada only a small group of "chartered" banks has this privilege. In most of the leading nations of Europe, such as Great Britain, France, Germany, etc., the business of note issue is virtually the monopoly of a single great central bank, more or less closely associated with the government. Legislation further puts a limit to the amount of notes which the banks may issue, prescribes certain forms of security and a certain amount of cash reserve, determines the form and denomination of notes, etc. As a result the issue of bank notes is in most countries a narrowly restricted and somewhat artificial function as compared with the business of deposit. Further details of the regulation of note issue will appear in connection with the study of banking systems in the next chapter.

Bank credit as a medium of exchange. From our previous study of the subject of money it appeared that the principal function of money was to serve as a medium of exchange. From our study of banking it is evident that the modern community makes constant use of bank credit for a medium of exchange. Now bank notes are money. But deposits are not, their acceptability in exchange for other things being not general but limited. Yet in the volume and importance of the service, bank deposits far outstrip bank notes.[1] In the United States for example the bank notes in circulation amount to almost two and a half billion dollars ($2,414,893,345 on December 1, 1925), whereas the amount of individual deposits subject to check in all the banks of the United States is over 15 billion dollars. The statistics of bank clearings presented in connection with a previous topic also furnish evidence regarding the importance of bank deposits as a medium of exchange. In the year 1924 the total clearings in 189 cities of the United States were over 455 billion dollars. It has been estimated that between eighty and ninety per cent of all exchanges are consummated by means of bank deposits.

---

[1] This statement is particularly applicable to the United States, Great Britain, and other Anglo-Saxon communities, where the deposit system with transfer by check is very highly developed. In the countries of Continental Europe there is relatively less reliance upon deposits and more upon bank notes.

Currency. It appears then that, while money exists primarily to serve as a medium of exchange, the great majority of exchanges are put through without the use of money. This makes it convenient to have another term, broader than money, which shall include everything that serves as a medium of exchange. The term *currency* has been adopted, and it may be useful at this point to present a classification of currency, embracing our previous classification of money (but omitting fiat money).

CURRENCY

I. Money.
    1. Standard money.
    2. Fiduciary money.
        (1) Representative money.
        (2) Token money.
        (3) Credit money.
            a. Government notes.
            b. Bank notes.
II. Bank deposits.
III. Other forms of currency (postal money orders, express checks, etc.).

In noting the great importance of bank deposits and other forms of currency which are not money, the reader will of course not lose sight of the fact that all such forms of currency are themselves rights to receive money; in other words, that the whole currency system is definitely based upon money, indeed upon standard money. We are not here concerned with that system in which fiat money has displaced standard money and thus become the basis of all the other forms of currency.

The need of elastic currency. A peculiar function of bank credit and one of the utmost importance is to furnish an element of elasticity to the currency system. No community's need of money is uniform at all times and seasons. In an agricultural community for example the volume of exchanges is especially great in the late summer and fall, on account of the business of harvesting and marketing the crops. Unusually large sums are being paid in wages to farm laborers, the farmers' receipts are at a maximum, and the local merchants find their sales swelling. There is need of considerably more currency than during other seasons. In a manufacturing town the weekly pay envelopes present a special need of money on Saturdays. The payment of salaries on the first of the month and the prevailing habit of settling accounts at the stores and elsewhere during the first few days of each month make a monthly variation in the need of currency. Corporations very commonly make their payments of dividends and interest quarterly, usually on the first of January, April, July, and October, which gives a quarterly swing to the fluctuating need of currency. There are also the cyclical fluctuations, the successive periods of business activity and business depression, of " good times and bad times." These examples should be sufficient to demonstrate that any community's need of currency is a constantly fluctuating quantity, the resultant of forces of the greatest variety and complexity. Perfectly to meet this need, the volume of currency should fluctuate correspondingly and automatically.

Most money is relatively inelastic. A glance at the foregoing classification of the ordinary currency system, recalling what has been learned about the several kinds of money, will suffice to convince the reader that no item in this list till we come to bank notes is so related to the state of business as automatically to expand and contract with the fluctuating need of currency.

Elasticity of bank credit. With bank credit (notes and deposits) the situation is different. Recalling our discussion of the business of discount and deposit, it will be appreciated that the demand for loans at the bank arises directly out of business transactions, either recently completed or contemplated for the immediate future. When business is active the banks are called upon for many and large loans. And in accordance with the relation between loans on the one side and notes and deposits on the other side of the account, as the banks respond to the demand for loans they increase their issue of notes and their deposits. Currency thus expands to meet business needs. On the other hand when business declines the customers of the banks have less need for currency. They pay off their loans at the banks faster than new loans are made. In paying their loans they either return to the banks the notes previously issued or they draw checks which reduce deposits. Currency contracts in response to declining need. Evidently there is here that close and automatic connection between volume of currency and business activity which is the mark of an elastic medium of exchange. Without such elastic element in its currency system, the modern community would suffer great inconvenience from alternating periods of stringency and plethora of currency. Bank credit thus performs a service of which the other parts of the currency system are generally incapable.

Influence of reserves. The elasticity of bank credit is limited by the necessity which is upon the banks to maintain adequate reserves. As we have seen, expansion of the banks' notes and deposits reduces their reserve ratio and inclines them to slow down the rate of expansion by raising the rate of discount. On the other hand, contraction of demand liabilities raises the reserve ratio, and the bankers seek to stimulate business by lowering the rate of discount. Banking reserves thus act, not to prevent the normal elasticity of bank credit, but to put on the brakes and so to prevent extremes of either expansion or contraction.

The note broker. It sometimes happens that a business man chooses to make use of an intermediary in securing the funds he wants and prefers to place his paper in the hands of a *broker* for sale to some bank or institution seeking short-term investments. He may have discounted all the commercial paper that his own bank will take and yet be in need of working capital.

One of the middlemen who has occupied a place of some prominence in the past is the *note broker*. He is a broker in the true sense of the word, in that his function is to bring buyer and seller together, and he receives a commission for so doing. The note broker still exists, but in recent years his place has largely been taken by the commercial paper houses.

The commercial paper houses. These are brokerage firms; they get a commission from the borrower for the services which they perform. They differ from the note broker in that they make loans directly to borrowers, whereas the note broker agrees to try to sell the commercial paper of his clients and in case he does not succeed in so doing returns the paper. But the commercial paper house has no intention of investing its own funds in the paper of its clients save for the time which is necessarily spent in selling it. It expects to sell and as a general rule succeeds in selling to banks, insurance companies, or other companies and individuals who have funds available for short-term investment. The paper which it sells rarely runs for more than six months.

It may develop that the house has made a miscalculation of the absorbing capacity of the market, and it may be required to carry some of its clients' notes. In order to insure itself and also to protect the interests of those to whom it sells commercial paper it conducts a very rigid investigation into the affairs of each company. When it sells a note to a bank it does not guarantee its payment, but it does guarantee its genuineness and recommend its purchase. Large banking establishments usually conduct an independent investigation themselves, but the small country bank has to rely on the investigation of the commercial paper house.

Relatively little capital of its own is required by the commercial paper house. As has been stated, it does not invest in this paper but expects to pass it on directly. Nevertheless each house must at any one time have a fairly large current investment in paper, and to carry this involves the temporary tying up of funds. These funds are secured by borrowing, at a lower rate of interest than that used to discount commercial paper of clients, from the large commercial banks on the promissory note of the house, secured by pledging the commercial paper of customers as collateral. Funds amounting to eighty or ninety per cent of the value of the collateral can be obtained in this way. If the house is forced by market conditions to carry some of the paper for a considerable time, it may make a profit by the difference in the rate of interest at which it borrows and the rate at which the customers' notes are discounted. There is of course the possibility of large losses, for the inability of the market to absorb the paper may be due to financial stringency, which will soon be evidenced by a rise in the interest rate, and to carry the paper the house may have to borrow money at a rate higher than the rate which was used in discounting it.

Such houses perform distinct services both to the bankers and to the business world. For the manufacturer or merchant a wider and probably cheaper market is provided for his promissory notes than is offered by the local bank. He can get funds when the local sources of supply are exhausted. The bank on the other hand gains, in that it has a greater range of industries from which to select the paper it wishes to buy. Frequently a bank in a small city would otherwise be limited to the paper of one type of industry or to that of a group of middlemen handling the same product. Depression in that industry will cause more suffering to the bank than if it spreads its risks (of non-payment) among different industries. It further enables it to select notes maturing at convenient times, making it possible therefore to change the nature of its investments from time to time.

Trust companies. Our attention has thus far been directed to the nature and functions of the commercial banks. This is by far the most important class of banking institutions, represented in the United States by the national banks and the state banks. Only a brief glance will be required by the other kinds of banks.

The trust companies are corporations whose original function was the care of their customers' investments and other financial affairs, the handling of trust funds, the administration of estates, etc. They naturally developed the business of keeping deposits subject to check and of making loans to their customers. Today the trust companies of the United States, in addition to their own peculiar functions, generally perform all the ordinary banking functions with the exception of note issue. To that extent the foregoing discussion applies to them equally with the strictly commercial banks. The differences between commercial banks and trust companies are steadily becoming less as time goes on.

Savings banks. These are institutions which accept deposits (usually in small amounts) and make investments for their depositors. They do not discount commercial paper, and their deposits are not the result of loans to their customers. Also their deposits are not legally payable on demand nor transferable by check in the ordinary way. The savings banks may require notice of withdrawal, of 30 or 60 days or some other period, though they frequently waive this right. Transfer of deposits requires certain formalities, such as the use of a special order and the presentation of the pass book. Savings deposits therefore do not serve as a means of making payments by check as in the case of the deposits of the commercial banks. The resources of the savings banks are invested, not in commercial paper, but in stocks and bonds, mortgages, and other investments of a safe and conservative character. The income thus obtained is devoted, after paying expenses, to the payment of interest to the depositors. The savings banks are really investment agencies rather than banks in the ordinary sense. Their customers are generally people in moderate circumstances, women and children, and those not engaged in active business. The total amount that may be deposited by any one person is sometimes limited, in order to restrict the savings banks' facilities to this class. For such persons the savings banks perform a useful service in gathering together small savings and investing them conservatively where they will yield a small but generally certain income. Wealthy persons and active business men have their own means of investing their funds and usually expect a higher rate of return. Recently there has been a movement among the commercial banks to enter the savings field, by establishing "savings departments." There is no corresponding tendency of the savings banks to take up commercial banking.

Private banks. Mere mention should here be made of the private banks, unincorporated banks conducted by individuals or partnerships and engaged primarily in the business of making investments, promoting corporations, underwriting security issues, etc. Some of them are to be counted among the most wealthy business concerns, and they occupy an especially important position in international banking and finance, but they do not play an important part in the ordinary commercial banking world.

The function of credit. We may now profitably turn back to our starting point, in order more clearly to visualize the important service rendered by credit, and particularly by bank credit. We have emphasized the importance of bank credit because it is fundamental in our economic organization, but we should not gain the impression that it occupies the field to the exclusion of commercial credit; that is, credit extended by the manufacturer to the wholesaler, or by the wholesaler to the retailer. In a great many transactions the banker is not called upon at all for an extension of credit. The manufacturer is able to take care of himself, or the purchaser from whom he gets his raw materials gives him time in which to work them up and pay for them.

We may in fact distinguish two points; namely, the importance of credit and the importance of a particular kind of credit. Credit, whether commercial or banking, is a sort of solvent factor in our organization, enabling us in a sense to make use of future goods at the present time. For society as a whole this cannot hold true of course but it can for a given individual, for he gets present goods in return for a promise to pay in the future, and the transaction is conditioned on the belief of the lender in the borrower's ability to produce goods or services at some time in the future.

The gain to a given individual from the credit system may seem of little importance or interest, but there is a social gain as well as an individual gain. Recalling for a moment the situation illustrated in our example at the beginning of this chapter we can readily see that, if a system of payment on the delivery of the goods had prevailed at the time the consumers found themselves unable to pay for their goods, the retailer would have ordered no new goods from the wholesaler. The wholesaler would have ceased ordering goods from the manufacturer, who would have been forced to stop manufacturing. In a short time the consumers perhaps would have found themselves again in a position to purchase goods on a cash basis, and renewed buying would have been passed on through the retailer and the wholesaler to the manufacturer. But how much more economical is that use of the productive resources which permits of a steady flow of goods instead of a jerky succession of activity and quiescence. Any business becomes somewhat disorganized if it has to shut down periodically, and disorganization is an expensive thing for society. Credit, which enables some to make present use of the surpluses of others, facilitates that smooth working of our productive apparatus which is essential to economy. It is one of the factors which promote efficiency in production and which therefore tend to keep the output of consumable goods at a high level.

The peculiar importance of bank credit. Bank credit performs a certain function in society better and more cheaply than other kinds of credit. The resources of any chain of producers are great, no doubt, but they are limited when compared with the resources of society as a whole, including the active leaders of business, those who are active in business as wage earners, and those who live on incomes derived from investments. The banks stand at the central point of the financial world and serve as reservoirs which attract the savings of all. They are in touch with more people and more lines of industry than a given firm can possibly be. If there is for the time being a surplus of funds in one line of industry, it can be made available for other industries. Without the bank as an intermediary how for example would the proprietor of a cotton mill in Massachusetts know that the owners of brass factories in Connecticut had funds which would be idle for six months? And if he did discover this fact, on what basis could they be made available to him? It might be done, but it would be difficult and costly. The bank stands ready to receive the funds of the Connecticut manufacturer and to pay him interest on a time deposit for six months; it can loan these directly to the cotton manufacturer in return for a promissory note of six months' duration. The thing is accomplished directly, simply, and at little expense.

Furthermore the bank serves as a selective factor in the direction of funds. If funds are needed more urgently in one industry than in another they are directed toward that industry, for the greater the need the higher the rate of interest which will be paid to secure a loan, and, other things being equal, a bank in making its loans will give preference to those willing to pay the higher rate of interest.

EXERCISES

1. Draw up a bank statement, using the following items: Stocks and bonds, $80,000; Undivided profits, $30,000; Reserve, $50,000; Notes, $75,000; Loans and discounts, $350,000; Real estate, etc., $100,000; Surplus, $50,000; Capital stock, $100,000; Deposits, $325,000.

2. Assume that a customer discounts a note for $5,000 for 3 months, the bank's rate of discount being 6 per cent, taking payment in cash. What changes would this transaction make in the bank's statement?

3. Suppose that the proceeds in exercise 2 were taken as a credit on the customer's deposit account. What changes would the transaction make on the bank's statement?

4. (*a*) What is the reserve ratio of the bank in exercise 1?
   (*b*) What is the bank's reserve ratio after the transaction of exercise 2?
   (*c*) What is the bank's reserve ratio after the transaction of exercise 3?

5. The bank of exercise 1, after discounting the note as in exercise 2, holds it one month and then rediscounts it with the Federal Reserve Bank, the rate of the latter bank being 5 per cent, and the proceeds taken in notes of the Federal Reserve Bank. Draw up a new statement showing the bank's condition after the operation. What is now the bank's reserve ratio?

Test starts here

CHAPTER XXIII

BANKING SYSTEMS

Banking principles and banking systems. Banking systems, like monetary systems, have evolved in the different nations in accordance with local traditions, habits, and political institutions. The foregoing chapter has treated of the fundamental nature and the general principles of banking, applicable regardless of the peculiarities of particular banking systems. We shall now find it interesting and instructive to inquire how these general principles are carried out in the banking organizations of some of the leading nations, giving particular attention naturally to the banking system of the United States.

Direct daily contact between the banking organization and the people, through the essential functions of discount and deposit, is everywhere maintained by local banks or branch banks, privately owned and privately managed. Over and above these local banks there is to be found in most of the nations a great central bank which, if not owned and controlled by government, is at least very closely associated with the national government. To this central bank is usually entrusted three important services: (1) the function of note issue, which in the development of banking has quite commonly been taken away from the local banks and made a monopoly of the central bank, (2) the business of rediscount for the local commercial banks, and (3) the concentration and management of the nation's banking reserves. In addition the central bank is the banker of the government, keeps the government money, manages the public debt, and handles the government's financial business. The central bank may or may not have ordinary dealings of discount and deposit with the general public. Thus we find at the head of the English banking system the Bank of England, in France the Bank of France, central government banks in Germany, Austria, Hungary, Italy, and most of the other European nations, the Federal Reserve System (with certain peculiar features) in the United States, and central government banks in many of the other countries of the Western hemisphere, to which Canada offers the most conspicuous exception.

The French banking system. This general account finds its simplest exemplification in France. The French public depends for its ordinary banking services mainly upon a few large banks located in the great cities and with branches scattered about the country. Small local banking corporations and private banks exist but are not numerous or especially flourishing. The Bank of France also deals directly with the public, though its requirement of three-name paper seriously limits its discounting operations. France is not particularly well provided with local facilities for discount and deposit.

The Bank of France. This is one of the oldest, largest, and most famous of central banks. It is owned by private capitalists, but the French government, though having no part in its ownership, has control over its management through the right of the President to appoint the Bank's governor and sub-governor, the former having direct charge of the bank's business and the appointment of all its employees. The Bank has the exclusive right of note issue. This function is of more importance in France than in the Anglo-Saxon countries, because of the relatively undeveloped state of deposit banking and the strong reliance upon bank credit in the form of notes. The notes of the bank of France are the principal element of the French monetary system. They are issued in connection with loans to the public, rediscount of commercial paper for the other banks, and loans to the government. They have also been given out in great volume in exchange for gold, making them to a certain extent what we have called representative money. The result is that, while no distinction is made between notes exchanged for gold and those issued in discount of commercial paper, the gold reserve back of the entire note issue has usually been very large.

While it has direct dealings with the public and for this purpose maintains branches and offices throughout the country, the Bank of France is primarily a bankers' bank. It holds deposits of the other banks, rediscounts for them, and its great hoard of gold is the national reserve reservoir, to be drawn upon as needed by the other banks through rediscount or by cashing their deposits.

The following condensed balance sheet will give further evidence as to the character of the business of the Bank of France.

STATEMENT OF THE BANK OF FRANCE, FEBRUARY 5, 1914 [1]

Liabilities	Francs	*Resources*	Francs
Capital	182,500,000	Cash	4,199,302,964
Surplus, etc.	68,627,487	Discounts	1,537,675,658
Notes	6,028,809,940	Advances on bullion and securities	803,806,112
Treasury account	145,501,708	Advances to the state	200,000,000
Deposits	712,384,776	Other resources	639,872,456
Other liabilities	242,833,279		
	7,380,657,190		7,380,657,190

British banking. The business of deposit banking is more highly developed in England than anywhere else in the world. A comparatively small number of large joint stock banks with thousands of branches, together with a smaller and declining number of private banks, bring the services of discount and deposit into every corner of the country. The use of deposits with payment by check is universal. Note issue is of relatively small importance, and the right of the joint stock banks to issue notes is negligible, being confined to those banks which were in existence in 1844 and limited to the amounts which those banks had outstanding at that time. The act of 1844 gave thus a partial monopoly of note issue to the Bank of England, which has now become almost complete.

The Bank of England. At the head of this banking system stands the Bank of England, founded in 1694, and with a long and illustrious history, out of which have developed peculiarities which set it apart in a class of its own. The Bank of England is in no sense a government bank, except that it is the fiscal agent of the British government. Its capital is all privately owned, and it is managed by directors chosen by the stockholders like any ordinary corporation. Though it handles all the British government's financial business, it has the least connection with government of any of the great central banks. While it deals directly with the public, through discount and deposit, it does not wage a vigorous competition with the other joint stock banks, in whose hands is the bulk of the nation's banking business.

---

[1] Owing to the extraordinary services of the Bank to the French government during the World War and the extraordinary inflation of its note issue then and since, a statement for a date before the beginning of the war gives a truer picture of the fundamental character of the Bank's business than would be obtained from a later statement.

Concentration of reserves. The Bank of England is content to be a central bank — a bankers' bank — and in this respect its position is supreme. Every British bank, including even those in Scotland and Ireland, keeps a part of its reserve in London, either in the Bank of England or in one of the great London banks. And the London banks keep practically all their reserves on deposit with the Bank of England, retaining only till money in their own possession. The Bank of England is thus the custodian of virtually the entire banking reserve of the United Kingdom, giving a degree of concentration which is to be found in no other country. Moreover the British banking system does business on a smaller percentage of cash reserve than is required anywhere else; it has been estimated that the deposit credit of Great Britain is based upon a money reserve of less than six per cent.

Note issue. It is in the system of note issue that the peculiar constitution of the Bank of England appears most strikingly. The system dates from the bank act of 1844, when the Bank was already 150 years old. By this act the Bank was separated into two distinct parts, called respectively the "Issue Department" and the "Banking Department." The latter conducts a regular banking business (except note issue) such as we have already described. The issue of notes is in the province of the Issue Department and is based upon a principle quite distinct from those which we have thus far studied and have seen illustrated in the Bank of France. Notes issued by the Bank of England must be secured either by British government obligations or by gold, and it was mainly to insure this that the Issue Department was separated from the rest of the Bank. In 1844 the government owed the Bank the sum of £11,015,100. These securities were put in the Issue Department as its assets, and it was allowed to issue notes to the same amount. Since then the Bank has been permitted to increase by about nine million pounds the amount of its notes secured by government obligations. All other notes issued by the Bank must be secured by gold coin or bullion, and to the amount so secured there is no limit. This arrangement will be made clear by a glance at the balance sheet below:

STATEMENT OF THE BANK OF ENGLAND, MARCH 4, 1925

Issue Department

Liabilities		*Assets*	
Notes issued	£146,530,345	Government debt	£ 11,015,100
		Other securities	8,734,900
		Gold coin and bullion	126,780,345
	£146,530,345		£146,530,345

Banking Department

Liabilities		*Assets*	
Proprietor's capital	£ 14,553,000	Govt. securities	£ 43,606,830
Rest	3,699,603	Other securities	80,048,938
Public deposits	9,825,224	Notes	21,703,580
Other deposits	119,104,913	Gold and silver coin	1,831,127
Seven-day and other bills	7,735		
	£147,190,475		£147,190,475

The first two items on the asset side of the Issue Department's statement are obligations of the British Government, the original sum held in 1844 being recorded separately from the later additions. Of the total note issue, £19,750,000 is therefore secured by government debt; the balance, £126,780,345, is really representative money, like the American gold certificates, though not distinguished in any way from the notes secured by government debt. In this respect similarity with the Bank of France will be noted; but here the similarity ends. In the Bank of France, as in all our discussion of general banking principles, no particular liability is matched against a particular item of assets. This is the ordinary situation with corporations in general. All the liabilities of a corporation are secured, as a whole, by all of its assets. Note the statement of the Bank of France or any of the illustrative bank statements which we have studied. But in the Bank of England one liability item; namely, notes issued, has been set apart, and for its security there have likewise been set apart two particular items of assets. The other liabilities of the Bank, as shown in the statement of the Banking Department, must find other assets to balance them.

Elasticity is sacrificed. Some very important results emerge from this peculiar system of note issue. The amount of notes which may be issued on the security of government debt is practically fixed. No expansion is possible here, and since these notes constitute a relatively small part of the nation's circulating medium, there is also no chance of contraction. These notes are simply a permanent and practically unvarying element in the monetary system. Changes in the volume of Bank of England notes can come only when gold is received by the Bank in exchange for its notes or *vice versa*. But since the gold might itself circulate as coin, such exchange means no real variation in the amount of money in circulation. In short the notes of the Bank of England do not furnish that elastic element in the currency system which we have seen it is one of the principal purposes of bank credit to provide. The notes of the Bank of England have been given perfect security but at the expense of elasticity.

How Great Britain secures elasticity. If it be asked how Great Britain manages with a currency system so seriously defective in this respect, the answer is twofold. In the first place, as we have seen, the British have come to rely upon bank deposits to an extraordinary degree, and the deposit system of Great Britain has well-nigh perfect elasticity. Secondly the gold element in the British monetary system is elastic to a degree found nowhere else in the world. London has long occupied a peculiar position as the financial and monetary centre of the world. Here is the world's greatest market for notes, drafts, bills of exchange, and all forms of commercial paper. A rise in the rate of interest in London causes a flow of capital from the ends of the world seeking the favorable opportunity for investment. A fall in the interest rate checks the inflow or even starts it in the opposite direction. The current interest rates in London are strongly influenced, if not controlled, by the rate of discount of the Bank of England. Thus the Bank of England holds the lever by which it controls the international flow of gold. A monetary stringency in England brings a rise in the interest rate and so an inflow of gold; an excess of currency causes the interest rate to fall, and some of the superfluous gold is drained off abroad.

Extreme procedure in time of crisis. There have been, it is true, certain occasions when the British credit system, functioning well enough in ordinary times, has broken down. In time of financial crisis there is need of great and sudden expansion of bank credit or, more correctly, there is need of expansion at some point to make good the panicky contraction of the regular sources of credit. Under such circumstances reliance is elsewhere upon the notes of the central bank. The Bank of England, with its rigid system of note issue, has on several occasions been unable to meet this demand. It has then been found necessary for Parliament to "suspend the bank act" by permitting an additional "uncovered" issue of notes; that is, of notes not secured by gold. This rather curious and clumsy device has never failed to produce the desired result of promising sufficient credit and allaying the panic.

The Banking Department. A few words only will be necessary to explain the second part of the Bank of England's statement as presented on page 450. The liabilities of the Banking Department include the capital stock ("proprietor's capital"), surplus ("rest"), deposits of the government ("public deposits"), and other deposits, a large part of which represents the Bank's obligation to the other banks. The "seven-day bills" are notes of the Bank which are not payable till seven days after their date, a form of bank note of which the British public makes some small use, especially on account of the reduced risk of loss when sent through the mails. On the asset side we find government securities, "other securities," which here means loans and discounts and all securities except government obligations, notes, and gold and silver coin. The notes are the notes of the Bank itself, which, in the statement of the Banking Department are assets; that is, they are claims of this Department against the Issue Department. The Banking Department can get notes at any time from the Issue Department in exchange for gold, and it can likewise always demand gold from the Issue Department to redeem bank notes which it may hold.

The reserve. By combining the two parts of the statement it will be seen that the Bank on this date had a net note liability of 125 million pounds, deposit liabilities of 129 millions, and a coin reserve of 128 millions or practically fifty per cent of its liabilities to other than its own stockholders. This large reserve ratio testifies to the Bank's position as the custodian of the nation's reserves and its preparedness to meet the demands of the other banks, the government, and the British public.

The Canadian banking system. As has been stated, a great central bank more or less closely related to the government stands at the head of the banking system in the majority of the leading nations of the world. There happens to be however on our northern border a nation which has maintained for a generation and more a highly efficient banking system without the aid of a central bank. A brief study of the Canadian banking system will fully reward our time and attention.

It was the bank act of 1890 which gave its essential characteristics to the present banking system of Canada. Banking is in the hands of a few large "chartered" banks, each of which maintains a great number of branches scattered all over the Dominion, thus bringing the facilities of one or more banks to the doors of practically every inhabitant. The business of discount and deposit is conducted in the usual way.

Notes. All of the banks issue notes, secured by their general assets in the normal manner, but having a prior lien ahead of other liabilities. The notes are further secured by a "safety fund" held and administered by the government, to which each bank contributes an amount of money equal to five per cent of its note issue. In case of failure of any bank to meet its note liabilities, the notes are paid out of this fund, which must then be brought back to five per cent of the total note circulation, either by sale of the assets of the failed bank or if necessary by further contributions from the other banks. Canada thus obtains well-nigh perfect safety for her bank note currency, without the earmarking of government bonds or any other particular assets of the banks as security.

The Canadian bank notes are also perfectly elastic. They are issued through the branches all over the country in connection with the discount of commercial paper. Each bank redeems on demand its own notes and also accepts the notes of the other banks, which are immediately presented to the bank of issue for redemption through the clearing houses in the large cities and through the nearest branches in the rural sections, in exactly the manner with which we have become familiar in the case of checks. The note circulation thus expands and contracts in response to business needs with the same facility as deposits and bears the same relation to bank reserves.

Few banks, but large ones. When the present system went into effect in 1890 it embraced 38 banks. These were mostly large institutions, their average paid-up capital being over $1,500,000, and the law provided that no new bank should be chartered with a capital of less than $500,000, of which at least half must be paid up. Since then the number of banks has steadily diminished, in part through failures and liquidations, but mainly on account of consolidations. In 1901, there were 34 banks, ten years later (1911) the number was 29, and the more rapid decline of recent years has brought the present number (September 30, 1925) down to 11. At the same time the size of the banks has been increasing, as well as the number of their branches. The average paid-up capital, which was over $1,500,000 in 1890, had become practically $2,000,000 in 1901, nearly $3,500,000 in 1911, and is at present over $10,000,000. And the eleven existing banks have surpluses exceeding their combined paid-up capital. The number of branches was 426 in 1890, about 1,650 in 1907, 1,841 in 1911, and 4,040 at present (December 31, 1924) besides 189 outside the Dominion. The total paid-up capital of the 11 existing banks is $116,164,660; their total resources are $2,836,701,398 (figures as of September 30, 1925).

The record of the Canadian banking system for safety and efficiency has been excellent. Failures have been few, with no losses whatever to note holders and only moderate losses to other creditors. Banking facilities have been provided and extended to meet the growth of the country with truly remarkable success. The lack of a central bank is compensated by the small number and great size and strength of the individual banks and by a certain supervision and control over all the banks exercised by an association of the banks themselves.

Colonial "banks." The limits of this book will not permit us to extend farther our investigation of foreign banking systems but require us now to turn to the study of the banking system of the United States. The so-called "banks" of the American colonial period were not real banks, but merely "batches of paper money" as one historian puts it; they did not engage in the business of discount and deposit and even in the issue of their notes did not conform to the principles of banking which our study has developed. During the Revolution and the years immediately following a few real banks were founded, which in addition to issuing notes made a start in the field of discount and deposit.

First Bank of the United States. In 1791, two years after the formation of the national government, Congress established a central bank, the first Bank of the United States. This bank, similar in its general make-up to the Bank of France, performed all the normal functions of a central bank, was the fiscal agent of the United States government, and did a regular banking business with the general public in competition with the growing company of local banks. Its notes became the principal circulating medium of the country. It was wisely and conservatively managed, performed its various functions well, and exercised also a salutary influence over the local banks, whose tendency to wild and unsound methods was apparent from the first. Indeed it was this check upon the ambitions of the local bankers and its competition with them that finally proved the Bank's undoing. Opposition arose, and Congress eventually failed to renew its twenty-year charter, which expired by limitation in 1811.

Local banks, 1811–1816. For the next five years the nation's banking was wholly in the hands of the local banks, chartered by the several states. These banks increased prodigiously in number. Relieved of the restraining influence of the great central bank, the majority ran wild in the paths of reckless and unsound banking. The currency was flooded with their notes, mostly of doubtful redeemability. They failed to perform satisfactorily the fiscal services which the Bank of the United States had regularly provided the government. Finally in the financial panic of 1814 the whole banking structure went to pieces, a condition made doubly disastrous as it came in the midst of the War of 1812.

The second Bank of the United States. The people learned the lesson, their misgivings against a strong central bank were overcome, and in 1816 the second Bank of the United States was founded. This bank's record is in its essential features a repetition of that of the first Bank, including its notable services to the government and the community and its final engulfment in the bog of political cross purposes. Its charter was not renewed, and it closed its doors in 1836. There was now a repetition of the experiences of 1811–1816, and the banking system collapsed in the panic of 1837.

Independent Treasury and local banks. Reëstablishment of a central bank proving politically impossible, the United States government cut itself loose from banking altogether by the establishment (first in 1844 and permanently in 1846) of the Independent Treasury for the custody of all government funds and the management of the government's financial business. The public was left to rely for banking services upon the local banks. Gradually some order came out of the chaos, and in the older and more settled parts of the country, particularly New England and New York, many sound banks arose and became more or less closely associated in an excellent banking system. "Wild-cat" banking continued to prevail in the frontier regions, and the country suffered grievously, especially from the multiplicity of note issues of all degrees of goodness and badness. Thus things continued down to the time of the Civil War.

The national banking system. The Civil War ushered in a new chapter in American banking history. Laws passed in 1863 and 1864 provided for a new class of banks chartered by the United States government, thus inaugurating the national banking system. The purpose of Congress in this legislation was primarily to aid the financing of the war and only secondarily to reform the country's banking system. In harmony with the first purpose, the law provided (as amended after the War) that each of the new national banks with a capital in excess of $150,000 must, before commencing business, purchase and deposit with the Treasurer of the United States bonds of the United States to the amount of at least $50,000 and that each national bank with a capital of $150,000 or less must deposit an amount of bonds equal to at least one fourth of its capital. Upon the security of the United States bonds so deposited the bank was permitted to issue its notes in any amount not in excess of the par value[1] of the bonds or the capital of the bank. National banks issue notes also secured by the deposit of lawful money. By an act of 1865 Congress imposed an annual tax of ten per cent upon the notes issued by state banks. This high tax made note issue unprofitable and reserved that function to the national banks.

There was also relaxation of the Independent Treasury system to the extent of permitting deposit of government funds in selected national banks and the receipt and disbursement of national bank notes by the Treasury.

The whole banking system. After the first few years the organization of the new national banks proceeded rapidly, and they soon assumed a position of predominance in the American banking system which they have since maintained. The national banks, the state-chartered banking institutions (state banks, trust companies, and savings banks), and the private banks have existed side by side, dividing among themselves the country's banking business, with only the function of note issue and the right to receive United States government deposits the exclusive prerogatives of the national banks. The following table presents a picture of the present American banking system, representing all banking institutions of which the United States Comptroller of the Currency had received reports as of June 30, 1924.

---

[1] Originally and up to 1900 the limit was 90 per cent of the par value or market value, whichever was smaller.

Kind of Bank	*Number of Banks*	*Paid-up Capital*	*Total Resources*
National banks	8,085	$1,334,011,000	$22,565,919,000
State (commercial) banks	17,436	1,061,619,000	14,816,011,000
Loan and trust companies	1,664	621,015,000	10,323,777,000
Stock savings banks	990	86,387,000	1,923,384,000
Mutual savings banks	613	....	7,364,656,000
Private banks	560	11,171,000	150,943,000
Total	29,348	$3,114,203,000	$57,144,690,000

It will be noted that, although the number of the national banks is less than half that of the state banks and trust companies, their aggregate capital and total resources are almost as great, owing to the greater average size and strength of the national banks.

Security for national bank notes. From what has already been stated it will be observed that the notes of the national banks are not issued freely upon the security of their general assets, in accordance with the simple banking principle as exemplified by the Bank of France and the Canadian banks, but are secured by a particular class of assets set aside for that purpose, upon the principle followed by the Bank of England. The United States law further provides for the safety of the national bank notes by imposing upon the United States Treasury the obligation to redeem any such note on demand, with of course the consequent right to collect the amount from the bank of issue. For this purpose each bank is required to maintain a cash fund with the United States Treasury equal to five per cent of its outstanding notes. The holder of a national bank note therefore has no concern as to the soundness of the bank of issue, relying as he does upon his claim against the United States; and the deposit of bonds is primarily to safeguard the Treasury rather than the note holder.

Elasticity destroyed. The national bank notes have thus been made eminently safe, but in so doing their elasticity has been sacrificed, just as in England. Since government bonds must be bought and deposited before additional notes may be issued, the banks cannot quickly expand their note issues in time of need. On the other hand, since the United States bonds bear a low rate of interest and cannot be profitably sold at short notice, the banks have in recent years generally owned little more than the minimum which the law requires and have sought to keep all their notes permanently in circulation in order to avoid loss on the investment. There is therefore no room for contraction of note issue to correspond with slackening business.

Disturbing fluctuations in volume. Elasticity in response to business needs is thus lost. On the other hand the volume of note issue has been made responsive to the condition of the public debt and the prices at which government bonds could be bought and sold. Thus have resulted some remarkable ups and downs, quite independent of the country's need of currency. From April 23, 1880, to October 2, 1890, the circulation of the national banks declined from $320,759,472 to $112,928,085, a decrease of sixty-two per cent, during a decade marked by large growth in the country's population and wealth and remarkable business expansion. There were significant gains in every other important item of the combined balance sheet of the national banks; individual deposits increased ninety-eight per cent and total resources fifty-nine per cent; only notes declined. The cause was the rapid retirement of the public debt, made possible by the great surplus revenues of that decade, reducing by more than one half the bonds available to secure circulation and sending skyward the prices of what remained. An opposite movement took place in the first decade of the twentieth century, when between June 30, 1899, and March 30, 1907, the circulation exactly trebled (rising from $199,358,383 to $597,212,063). The principal cause was new legislation in

1900, which affected the public debt and the profitableness of note issue. The national bank notes have thus not only failed to aid the country's business by a normal elasticity but have been a disturbing influence on account of the wide and irresponsible fluctuations in volume to which they have been subject.

Artificial value of government bonds. The Currency Act of 1900, to which reference is made in the previous paragraph, provided among other things for the ultimate refunding of the existing United States debt into a class of bonds paying two per cent interest and known as the "Consols of 1930." During the following decade refunding proceeded rapidly, and on May 31, 1911, four fifths of the total interest-bearing debt was in the two per cent bonds, either the Consols of 1930 or the new Panama Canal bonds. Now no government in the world could borrow at the rate of two per cent, without some artificial advantage. The advantage enjoyed by the United States was the utility of its bonds to the national banks to secure their circulation and the legal compulsion upon all national banks to own bonds equal in general to one fourth of their capital. By this time the debt of the United States had been so reduced that there were barely enough bonds to meet the requirements of the national banks. On May 31, 1911, the situation was thus:

INTEREST-BEARING DEBT OF THE UNITED STATES, MAY 31, 1911

	Two Per Cent Bonds	*Other Bonds*	*Total*
Total interest-bearing debt	$730,882,130	$182,435,360	$913,317,490
Deposited by national banks to secure circulation	659,617,720	37,823,580	697,441,300
Percentage of total so deposited	90%	21%	76%

The national banks were using three fourths of the entire interest-bearing debt of the United States to secure their note issues; of the two per cent bonds they had nine tenths. This artificial position of the two per cents has continued down to the present. In 1911, the defects of the bond-secured notes having become evident, the United States adopted the policy of making its future loans by means of bonds and notes not available for bank note security, and the national banks have ever since been limited to those classes of bonds only which were in existence on May 31, 1911. Since then all of these bonds except the two per cents have been retired, and there are now (December 31, 1925) available for securing circulation only 675 millions of two per cent bonds, of which the national banks have acquired for securing circulation 665 millions, or nearly ninety-nine per cent.[1]

Here then is an artificial market, which has regularly absorbed nearly the whole stock of United States two per cent bonds and given them a value far above their worth as an investment only. For the past year or two they have been quoted slightly above par and generally higher than other United States bonds paying three or four per cent. Evidently the circulation privilege is worth one per cent or more.

The illusion of bond security. Now the conclusion from these facts is that the much vaunted government bond security of the national bank notes is largely an illusion. Of course the security is sufficient for an isolated failure now and then. But any crisis which would threaten the whole bank note system would demonstrate that the two per cent bonds could not bring more than sixty-six cents on the dollar if sold on an investment basis. In short the notes are safe because secured by government bonds, and the bonds are valuable because they secure the notes. Such an arrangement is popularly known as "lifting yourself by your boot-straps." What actually gives the national bank notes their unquestioned safety is the unconditional obligation of the United States Treasury to redeem them.

A nation of small local banks. In addition to the lack of an elastic bank note currency, there has been in the United States until 1913 the absence of any system of centralized banking reserves, such as is provided in most countries by the central bank and in Canada by a small group of large and powerful banks with their far-reaching system of branches. The United States has been the country *par excellence* of small independent local banks.

---

[1] We have disregarded a few thousands ($39,300) of the matured four per cent loan of 1925 which were still on deposit to secure circulation.

Branch banking is forbidden by law to the national banks and not highly developed by the state banks and trust companies. A glance at the table on page 458 will show that (omitting the savings banks and private banks) the nation's commercial banking business is in the hands of more than 27,000 local banking institutions, with an average paid-up capital of only $115,000, and average resources of $1,767,000. Compare this with the situation in Canada, where there are only eleven banks, with an average paid-up capital of over $10,000,000 and total resources averaging more than $250,000,000.

The need of elastic currency and centralized reserves. Owing to the great extent of our territory, the diversity of our industries, and the predominant position of agriculture, the need of an elastic bank note currency and a system of centralized reserves, to which general reference has already been made, is particularly urgent in the United States. Until the foundation of the Federal Reserve system in 1913, this need was not met. Struggling to cope with the situation as best they might, the banks in the country districts and smaller towns and cities developed the custom of depositing as much of their resources as they felt they could spare in periods of slack business with banks in the large cities of their respective regions. These city banks in their turn made deposits in other larger cities. Thus were set up, as it were, little rivulets of currency springing from sources all over the country, uniting in the cities to form larger streams, joining again into still larger, till eventually a mighty flow of money poured into the banks of a few financial centres, the greatest of all being New York City. This flow of deposits was encouraged by the willingness of the large city banks to pay interest at a moderate rate upon deposits of other banks. The funds so obtained were loaned by the city banks to their customers, a considerable part being used for stock market speculation. On the other hand when a succeeding period of business activity opened up larger opportunities for the outlying banks or when a crisis threatened, they demanded the return of their reserves and the flow of currency was reversed. This forced the city banks to call in their loans and contract their deposits and caused a state of monetary stringency in these cities. This ebb and flow, besides its minor variations, had a great annual cycle on account of the need of currency for harvesting and marketing the crops. In the spring there was a period of easy money, a flow of currency to New York and a few other centres, low interest rates and encouragement to speculation in these cities. In the late summer and fall money flowed back to the country, while in the centres monetary stringency prevailed and interest rates soared.

It will be seen that, in a sense, the great city banks, particularly in New York, undertook to perform the services of a central reserve reservoir. But they were fatally handicapped by their great number, their independence of each other, their lack of any branch organization, and their inability to expand and contract their credit in the form of notes. Moreover they were unwilling to accept the responsibility of their position to the extent of keeping adequate reserves. The result was inevitable. Alternating periods of surplus currency and monetary stringency were the normal thing; inability of the outlying banks to obtain the return of their reserves in time to avoid disaster was common; bank failures and business failures were all too numerous; financial panics and crises were made more frequent and more severe. The situation grew steadily more intolerable with the development of the country's wealth and industry, until at last a remedy was found in the federal reserve system which became law on December 23, 1913.

The federal reserve system: organization. The federal reserve system furnishes the United States a centralized banking organization, performing in general the functions of a central bank which have been previously outlined and illustrated by reference to the Bank of France and the Bank of England. The American system is, however, peculiar in many respects. The essential features of its structure are as follows.

The whole country is divided into twelve "federal reserve districts," with a "federal reserve bank" located in a certain city in each district. The districts are distinguished by numbers, the following being a list of the districts and their respective central reserve cities: 1, Boston; 2, New York; 3, Philadelphia; 4, Cleveland; 5, Richmond; 6, Atlanta; 7, Chicago; 8, St. Louis; 9, Minneapolis; 10, Kansas City; 11, Dallas; 12, San Francisco. Each federal reserve bank is permitted to establish branches and agencies located in cities in its district. At present (February, 1926) there are twenty-three branches and three agencies. The accompanying map[1] shows the boundaries of the several districts and the location of the federal reserve banks and their branches and agencies.

Each federal reserve bank is a corporation, and its stock is all owned by the individual banks within its district which have come into the system. These banks are known as "member banks." Every national bank is required by law to be a member, but membership is optional for the state banks and trust companies. On January 31, 1926, there were 9,459 member banks of which 8,030 were national banks and 1,429 were state banks and trust companies. About one twelfth of the state banks and trust companies are members, but this twelfth represents more than half of the total resources of all the state banks and trust companies, showing that the largest and strongest institutions have pretty generally come in.

At the head of the whole system stands the Federal Reserve Board, composed of the Secretary of the Treasury, the Comptroller of the Currency, and six other members chosen by the President of the United States. Its office is in Washington. Each federal reserve bank is managed by a board of nine directors, of whom six are chosen by its member banks and three by the Federal Reserve Board.

Functions. The banking business of the federal reserve system is conducted by the twelve federal reserve banks. They rediscount commercial paper for their member banks, they hold deposits of the member banks, they issue notes, and they act to a certain extent as fiscal agents of the United States government. They are permitted also certain dealings with the general public, but their principal business is with their respective member banks,

---

[1] Two of the agencies are located in Havana, Cuba.

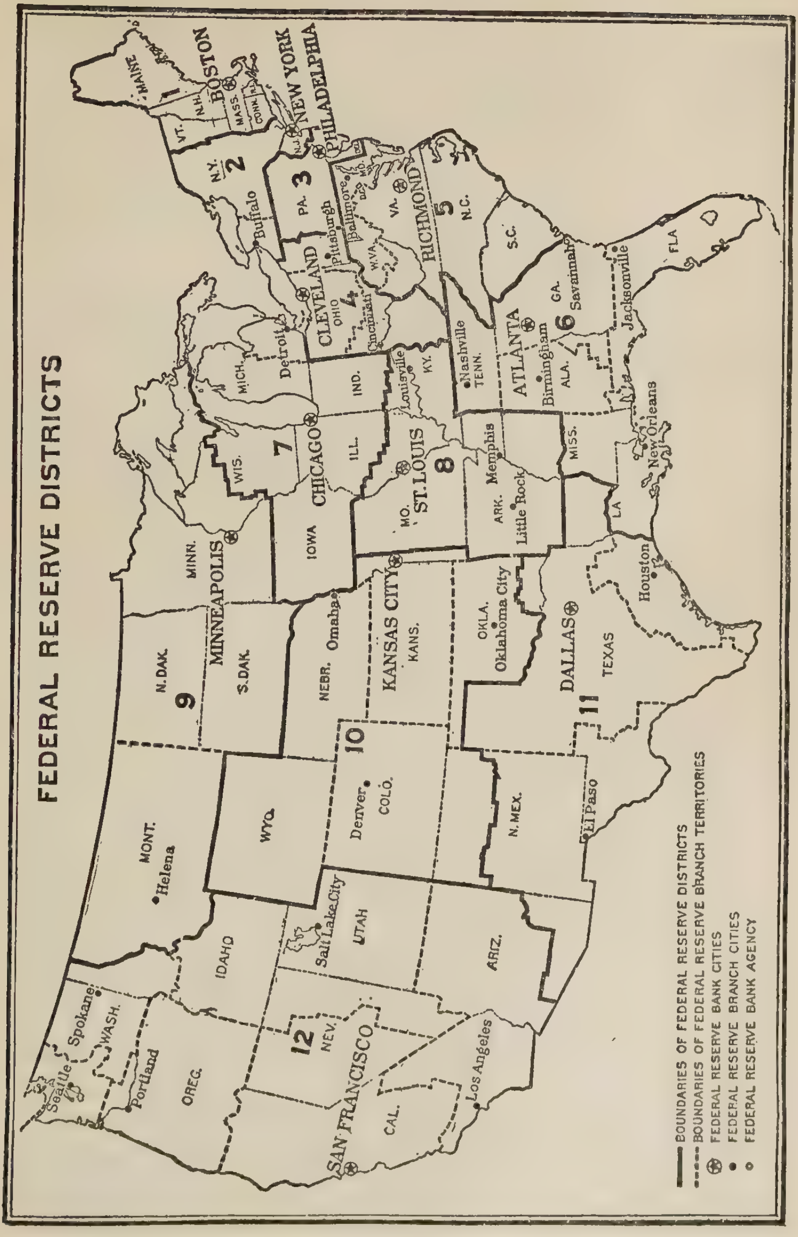

FEDERAL RESERVE DISTRICTS
1
BOSTON
MAINE
N.H.
VT.
MASS.
CONN.
R.I.
2
NEW YORK
N.Y.
N.J.
Buffalo
3
PHILADELPHIA
PA.
DEL.
4
CLEVELAND
OHIO
Pittsburgh
Cincinnati
5
RICHMOND
VA.
W.VA.
MD.
D.C.
Baltimore
N.C.
S.C.
6
ATLANTA
GA.
FLA.
ALA.
TENN.
Nashville
Birmingham
Savannah
Jacksonville
New Orleans
7
CHICAGO
MICH.
Detroit
IND.
ILL.
WIS.
IOWA
8
ST. LOUIS
MO.
KY.
Louisville
ARK.
Memphis
Little Rock
MISS.
9
MINNEAPOLIS
MINN.
N.DAK.
S.DAK.
MONT.
Helena
10
KANSAS CITY
KANS.
NEBR.
Omaha
COLO.
Denver
WYO.
OKLA.
Oklahoma City
11
DALLAS
TEXAS
Houston
El Paso
LA.
N.MEX.
12
SAN FRANCISCO
CAL.
Los Angeles
NEV.
OREG.
Portland
WASH.
Seattle
Spokane
IDAHO
UTAH
Salt Lake City
ARIZ.
BOUNDARIES OF FEDERAL RESERVE DISTRICTS
BOUNDARIES OF FEDERAL RESERVE BRANCH TERRITORIES
FEDERAL RESERVE BANK CITIES
FEDERAL RESERVE BRANCH CITIES
FEDERAL RESERVE BANK AGENCY

Fig. 42.

with the government, and with each other. Although it might appear that instead of a centralized banking system for the nation we had twelve central banks each in its own district, the fact is that the Federal Reserve Board, through its appointment of three directors of each district bank and the other powers over their banking business which the law gives it, exercises a very real control and makes the system, in spite of its peculiar district organization, a real central banking institution.

Rediscounting and discounting. The rediscounting of the federal reserve banks is in harmony with the general principles with which the reader is now familiar. The law makes certain restrictions as to the kind of commercial paper that may be rediscounted, prescribing that it must be of short duration, based upon commercial transactions or upon loans on collateral of United States government obligations, and bearing the endorsement of a member bank. Any member bank is thus enabled to strengthen its position when necessary by exchanging commercial paper either for deposits with the federal reserve bank or for its notes. The discount dealings with the public, called " open-market " operations, consists of buying and selling certain specified forms of commercial paper, as is done by any commercial bank. The reserve banks also have dealings in commercial paper with one another.[1] Through the discount rates of the reserve banks, the federal reserve system has an important influence upon the nation's discount market, but it has never approached in this respect the predominant position held by the Bank of England, the Bank of France, and other great central banks.

Deposits. The federal reserve banks receive ordinary deposits only from their member banks and the United States government, but deposits for the purposes of exchange and collection are received also from other federal reserve banks and from state banks

---

[1] As a result of their rediscounts and open market operations the twelve federal reserve banks held on December 16, 1925:

Bills discounted, secured by government obligations	$343,121,000
Other bills discounted	275,946,000
Total bills discounted	$619,067,000
Bills bought in open market	$352,692,000

and trust companies which are not members of the system. Against its deposits each federal reserve bank is required normally to keep a reserve of gold or lawful money of at least thirty-five per cent, though a smaller ratio may be permitted subject to a progressive tax on the deficiency.

Each member bank is required to keep with the reserve bank of its district a deposit not less than a certain percentage of its own deposits; namely, three per cent for time deposits, and either seven, ten, or thirteen per cent for demand deposits, depending upon the size of the city in which the member bank is located. This is called the bank's "lawful reserve" and supersedes a complicated requirement of minimum reserves against deposits formerly prescribed by the national banking law.[1]

Federal reserve notes. Two kinds of notes are issued by the federal reserve banks. The *federal reserve notes*, though nominally the obligations of the United States government, are essentially bank notes of the normal type, secured not by government bonds but by commercial paper and gold. The law requires each federal reserve bank to maintain a gold reserve against its notes, which must normally be at least forty per cent but which may on occasion fall below that limit, subject to a progressive tax upon the amount of the deficiency. These notes, issued in the process of the bank's rediscounting and open market operations or in payment of its deposit liabilities, are thus capable of furnishing an elastic currency of unquestioned strength. Their amount rose rapidly to 547 millions on June 30, 1917. The federal reserve system was then drafted into the service of the war, resulting in extraordinary issues of federal reserve notes, whose amount outstanding was nearly three and a half billions on June 30, 1920. Since then there has been a material decline, but the federal reserve notes today still constitute the largest single item in the monetary circulation;

---

[1] On December 16, 1925, the combined balance sheet of the twelve federal reserve banks showed the following deposits:

Member bank reserve account	$2,264,797,000
United States government deposits	5,954,000
Foreign bank deposits	8,398,000
Other deposits	21,356,000
Total deposits	$2,300,505,000

the amount outstanding on December 1, 1925, being $2,043,940,055 of which $1,741,964,886 were in circulation.

Plan to retire the national bank notes. It was evidently the intent of the authors of the Federal Reserve Act that the notes of the federal reserve banks should ultimately displace at least the greater part of the existing national bank notes, whose glaring defects furnished one of the motives for the establishment of the new banking system. The law therefore repealed that part of the national banking act which required each national bank to invest a certain part of its capital in United States bonds. It further provided that the national banks might sell to the federal reserve banks at par bonds which they had deposited to secure their notes and might retire the corresponding notes. This privilege was limited to a twenty-year period beginning two years after the enactment of the law (1913) and to a maximum amount of $25,000,000 a year. The federal reserve banks, upon purchase of such bonds, were permitted to issue their own notes secured by the bonds. These notes are called *federal reserve bank notes* and are in all essential respects like the national bank notes which they replaced. As an alternative to issuing federal reserve bank notes, the federal reserve banks were permitted to exchange the two per cent bonds, thus bought from the banks, with the United States Treasury for new three per cent bonds or notes without the circulation privilege.

How the plan has worked. This program fully carried out, assuming the federal reserve banks adopted the alternative of converting the two per cent bonds instead of issuing notes upon their security, would have diminished the bond-secured bank notes and also the United States two per cent debt at the rate of $25,000,000 a year beginning in 1916. For about a year and a quarter things went according to plan. By June 30, 1917, the two per cent debt of the United States had been reduced from 731 millions to 675 millions (not all however due to this particular cause). The amount of national bank notes secured by United States bonds declined from 720 millions on December 31, 1915, to 668 millions on June 30, 1917.

Further progress however was stopped by the World War, during and following which there was an actual increase in the bond-secured notes, which brought their amount up to nearly 746 millions in April, 1924, exceeding the amount in 1915 by 26 millions. Since then there has been a decline to 663 millions (on November 30, 1925),[1] but this was primarily due to the retirement in February, 1925, of the remaining 63 millions of the four per cent loan of 1925, the last remaining issue, other than the two per cents, available to secure circulation. Recently the two per cent bonds have stood slightly above par on the market, making the banks disinclined to sell their holdings at par.

No further retirement of two per cent bonds has occurred, the amount outstanding at the close of the year 1925 being the same as in 1917; *i.e.*, 675 millions. The federal reserve bank notes have never become an important part of the monetary system. On December first, 1925, there were only 6½ millions outstanding, for whose retirement lawful money had been deposited with the Treasury by the federal reserve banks. It is evident therefore that the Federal Reserve Act has accomplished little toward the elimination of the national bank notes, and it appears hardly likely that it will be effective to accomplish much more in the near future. Further legislation will probably be required to bring about this result.

Centralized reserves. It will be apparent from the foregoing account of the business of the federal reserve banks that the system is well adapted to furnish a central reservoir and centralized control for the country's banking reserves. The only reserves against deposits now required by law of the member banks are their deposits with the federal reserve banks. Of course the individual banks must keep some cash reserves, in order to meet their daily obligations, but this is left to the discretion of the particular bank, and the amount actually kept is extremely small. For example, on September 28, 1925, the 8,085 national banks had together 362 million dollars "cash in vault" against "demand deposits" of 10,428 millions and total deposits of 19,930 millions. This was not quite three and one half per cent of demand deposits and less than two per cent of total deposits. Their "lawful reserves" on deposit with the federal reserve banks on this date were 1,324 millions, about thirteen per cent of their demand deposits and six and two thirds per cent of total deposits.

---

[1] The total amount of national bank notes outstanding is greater than this owing to the presence of notes secured by lawful money, of which there were on November 30, 1925, 48 millions, making the total of national bank notes outstanding on that date not quite 711 millions.

The federal reserve banks have on the other hand maintained very large reserves, generally far in excess of the legal requirements. Their gold holdings have almost from the start been extraordinarily swelled as a result of the suspension of gold payments throughout most of the world, the enormous demand for American exports, and other events of the World War; and these causes have not entirely ceased to operate down to the present time. On December 16, 1925, the gold holdings of the twelve federal reserve banks were 2,702 million dollars, against total liabilities for notes and deposits of four billions, showing a reserve ratio of sixty-six and three tenths per cent.[1]

Any member bank may always meet an additional need for currency by translating its deposit with the reserve bank into federal reserve notes. Moreover, by the process of rediscount, the member bank in sound condition may always increase its deposit with the reserve bank or its stock of federal reserve notes. Finally, through dealings between the reserve banks themselves directed and controlled by the Federal Reserve Board, it is possible to direct the flow of reserves to the particular districts where for the time being the need is urgent.

The reader will realize how elasticity of currency has been made effective through the machinery of rediscount, coupled with the power of the federal reserve banks to issue an elastic note currency. The local banks are now able to meet the needs of their customers for currency in times of special business activity, through granting deposits or giving federal reserve notes, and when the special need is over the currency contracts as loans are paid off, deposits decline, and federal reserve notes are returned to the federal reserve banks in exchange for commercial paper previously rediscounted and now paid off.

---

[1] For the details of the gold holdings of the twelve federal reserve banks on December 16, 1925, *cf.* the table on page 473 below.

As a result we now have in the United States for the first time in our history an effective mobilization of reserves and a truly elastic currency. Whereas in former days perfectly sound banks were sometimes forced to fail because they could not obtain in time the money due them from other banks, such an event is now virtually inconceivable. A local panic can no longer wreck a bank or a series of banks by a wild unreasoning "run" of depositors to withdraw their money. Any soundly managed bank can now through rediscount convert its other resources into cash sufficient to meet any run. Finally those nation-wide crises which from time to time sweep over the country have been largely shorn of their power to cause wreck and ruin to banks and business houses from the mere inability to convert sound assets into means of immediate payment. The power of the banks, aided by the federal reserve system, to meet the legitimate credit needs of their solvent customers is almost unlimited, and the banks can now come to the aid of the business community in time of trouble without jeopardizing their own safety. The old time financial panic, in which every business concern and every bank was engaged in a wild scramble to save its own skin, is probably forever a thing of the past.

A national system of clearing. The federal reserve system has also become a national clearing house for settling the claims of the reserve banks against each other and for collection of checks, drafts, and other claims of the individual banks against out of town banks. This mechanism is far more effective than anything that ever existed in the country before. It secures collection at par without charge in the shortest possible time and with a remarkably small transfer of actual money, thus doing away with what was formerly a cause of much delay and heavy expense to the public.

The government's financial business. The reader will recall that, following the banking collapse in 1837, Congress established the Independent Treasury system, whereby the United States government undertook to keep its own moneys and handle all its financial business without the aid of the banks. It will also be recalled that with the later establishment of the national banking system the Independent Treasury law was relaxed to the extent of permitting deposit of government funds in selected national banks. The government has thus conducted its financial business through the Treasury at Washington, sub-treasuries located in nine principal cities, and the offices of the mint, with the aid of a variable number of national bank depositories. The system has always been an anomaly, unlike anything to be found among other important nations, and fraught with danger to the country's monetary stability. Use of national bank depositories, chosen from the thousands of such banks, has been inconvenient, wasteful, and politically insidious.

The Federal Reserve Act accordingly provided that the funds of the United States (with certain exceptions) might be deposited with the federal reserve banks and that these banks might be required to act as fiscal agents for the Treasury. The Treasury was at first slow to avail itself of these facilities, but the exigencies of the World War, in particular the conduct of the government's loans and the handling of the funds so obtained, soon compelled the utmost reliance upon the federal reserve system, whose chief activities became concentrated upon its function as fiscal agent of the United States. At the same time the number of bank depositories was also greatly increased, and many " special depositories " are still used in connection with the government's borrowing operations. In 1921, under authority of a law of 1920, the sub-treasuries were abolished as no longer required, and the Treasury is making progress, slowly to be sure, toward curtailing its reliance upon the local banks and concentrating its fiscal business in the hands of the federal reserve banks. As illustrating the present situation, it may be noted that on December 17, 1925, the assets in the " general fund " of the United States Treasury amounted to 631 million dollars, of which 34 millions were on deposit in federal reserve banks, 329 millions in " special depositories," 28 millions in national banks, 2 millions deposited abroad, and the balance, 243 millions, in the custody of the Treasury itself.[1] The federal reserve system has still greater possibilities for service to the United States Treasury than have yet been availed of.

Statement of the federal reserve banks. The following is the combined statement of the twelve federal reserve banks on December 16, 1925:

Resources

Gold with federal reserve agents	$1,394,759,000
Gold redemption fund with U. S. Treasury	54,570,000
Gold held exclusively against federal reserve notes	$1,449,329,000
Gold settlement fund with Federal Reserve Board	664,899,000
Gold and gold certificates held by banks	587,358,000
Total gold reserve	$2,701,586,000
Reserves other than gold	108,358,000
Total reserves	$2,809,944,000
Non-reserve cash	45,663,000
Bills discounted, secured by government obligations	343,121,000
Other bills discounted	275,946,000
Total bills discounted	$619,067,000
Bills bought in open market	352,692,000
U. S. government securities, Bonds	73,451,000
Treasury notes	153,740,000
Certificates of indebtedness	171,280,000
Total U. S. govt. securities	$398,471,000
Other securities	3,195,000
Foreign loans on gold	8,798,000
Total bills and securities	$1,382,223,000
Due from foreign banks	710,000
Uncollected items	952,147,000
Bank promises	61,607,000
All other resources	71,632,000
Total resources	$5,269,926,000

Liabilities

Federal reserve notes in actual circulation	$1,788,230,000
Deposits, Member banks, reserve account	2,264,797,000
Government	5,954,000
Foreign bank	8,398,000
Other deposits	21,356,000
Total deposits	$2,300,505,000
Deferred availability items	827,072,000
Capital paid in	116,964,000
Surplus	217,837,000
All other liabilities	19,318,000
Total liabilities	$5,269,926,000

---

[1] The Treasury also keeps in its own custody the gold coin and bullion back of the gold certificates, the gold fund of the Federal Reserve Board, the gold reserve against the greenbacks, and the silver dollars back of the silver certificates and the Treasury notes of 1890.

The United States monetary system. The point has now been reached where it is possible to obtain a view of the whole monetary system of the United States, the separate parts of which have held our attention from time to time during the past four chapters. The people of the United States today make use of ten kinds of money; namely, three kinds of coin (standard, "hybrid," and token), two forms of representative money, and five forms of credit money. The table on the next page contains the list and shows also the location of each form of money as of December 1, 1925.

The present monetary system is not the result of a logical plan but the product of a long historical development. If the monetary system were being now constructed *de novo,* there would be no need of all these ten kinds of money. The silver dollars serve no useful purpose. They might be abolished or, if there is some real desire for them as pocket money in certain parts of the country, they could perfectly well be transformed into token coins by making them specifically redeemable in gold and limiting their legal tender quality. Whatever might become of the silver dollars, there is no advantage in maintaining the silver certificates. The greenbacks have been a source of embarrassment to the Treasury in the past, they are a possible element of danger for the future, and they serve no purpose not better met by other kinds of money. The monetary system would be improved if the United States should pay them off and destroy them. The Treasury notes of 1890, the national bank notes, and the federal reserve bank notes are all historical relics, anomalous forms of credit money whose usefulness is now superseded by the federal reserve notes. Supposing the changes thus suggested to have been made we should have a simplified monetary system consisting of just four kinds of money; namely, (1) gold coin and bullion, (2) token coins of silver and other metals, (3) gold certificates, and (4) federal reserve notes. This monetary system would have all the necessary elements represented by our present ten varieties of money. While some of these changes will come slowly if at all, it is evident that our monetary development is in the direction of some such system as is thus outlined.

CIRCULATION STATEMENT OF UNITED STATES MONEY, DECEMBER 1, 1925

Kind of money	*Held in U. S. Treasury*	*Held outside U. S. Treasury*			*Total stock*
		By federal reserve banks and agents	*In circulation*		
			Amount	*Per capita*	
1. Gold coin and bullion	$3,695,492,642	$ 304,952,935	$ 425,853,229	$ 3.71	$4,426,298,806
2. Silver dollars	454,683,558	14,302,707	54,684,813	.48	523,671,078
3. Token coins:					
a. Subsidiary silver	6,376,833	8,656,689	272,216,935	2.37	287,250,457
b. Minor coins (estimated)	3,000,000	6,000,000	100,000,000	.88	109,000,000
Total coin and bullion	$4,159,553,033	$ 333,912,331	$ 852,754,977	$ 7.44	$5,346,220,341
4. Gold certificates	....	$ 582,410,020	1,108,743,229	9.67	1,691,153,249*
5. Silver certificates	....	56,397,703	388,012,326	3.38	444,410,029*
Total certificates	....	$ 638,807,723	$1,496,755,555	$13.05	$2,135,563,278*
6. U. S. notes	2,854,129	$ 37,841,320	$ 305,985,567	$ 2.67	$ 346,681,016
7. Treas. notes of 1890	....	....	1,375,504	.01	1,375,504*
8. National bank notes	14,451,292	29,555,510	666,743,642	5.81	710,750,444
9. Federal reserve bank notes	195,979	97,517	6,184,817	.06	6,478,313
10. Federal reserve notes	1,000,755	300,974,414	1,741,964,886	15.19	2,043,940,055
Total notes	18,502,155	$ 368,468,761	$2,722,254,416	$23.74	$3,109,225,332
Total monetary stock	$4,178,055,188	$1,341,188,815	$5,071,764,948	$44.33	$8,454,070,169

* Not included in the total of this column, in order to avoid double counting.

CHAPTER XXIV

MONEY AND PRICES. THE VALUE OF MONEY

Value of money not constant. Thus far in our investigation of the problems of value and price the value of money has been considered a constant. This was a perfectly proper procedure, since value and price are not absolute quantities but relations between commodities and services arising from exchange. The value of anything is the quantity of some other thing that would be given in exchange for it. The price of anything is the quantity of money that would be given in exchange for one unit of it.[1] In studying these relations of other things to money, it has been convenient and permissible to consider the value of money itself constant and the changes in the relations between money and other things as changes in the values of the other things.

Such *a priori* assumption is of course quite arbitrary. If one bushel of wheat exchanged for two bushels of corn yesterday and for three bushels of corn today, we may say that the value of a bushel of wheat has *risen*, from two bushels of corn to three bushels, assuming thus that the value of corn is a constant. But we could just as correctly assume the value of wheat constant and say that the value of a bushel of corn has *fallen*, from one half bushel of wheat to one third bushel. Without knowledge of the facts of the particular case, the change in relation may have been due to a change in either corn or wheat alone or to changes in both of them. If two cars happen to be at the same level in an elevator shaft and later it is found that car A is above car B, the change in relation may have been caused by (1) a movement upward of A while B stood still, (2) a downward movement of B while A stood still, (3) an upward movement of A and a downward movement of B, (4) upward movements of both, A moving farther than B, or

---

[1] See Chapter I.

(5) downward movements of both, B moving farther than A. And, to complete the case, both cars might move the same distance either up or down, leaving their relative positions unchanged. So it is of the value relation between any two commodities.

Nor is this conclusion altered if one of the commodities is that one which has been accepted as the money of the community, as a result of which it has become the custom to express the values and prices of all other commodities and services in terms of this one commodity. Gold is the commodity which has been generally so accepted by the modern world. But gold, although it is more stable in value than some other things, has still no qualities which put it above the laws of demand and supply and release it from the universal tendency to change in value. And the adoption of gold as the money commodity had no power to give a fixed and immutable state to that which was theretofore subject to change. Under a money economy value and price are relations between money and other things, and a change in the price or value of any other commodity may be the result of changes either in itself or in the money commodity. In short it is always possible that a change in the value of anything may be due to a change in the value of money.

What is the value of money? This brings us to the necessity of a workable definition of the value of money, in the search for which we shall find it convenient to depart slightly from the analogy of the definition of value in general. The value of any commodity may be expressed in terms of any other commodity, but this is clumsy and confusing, and by common consent we have come to express all values in terms of money. In the same way the value of money might be expressed in terms of any other commodity. This would likewise be clumsy and unprofitable, but there is not here the same escape. We cannot obviously express the value of money in terms of itself. Moreover the concept of the value of money, to be useful, must not be its relation to some one commodity. What we are interested in is the relation of money to all other commodities and services, to economic goods in general. When we say money has depreciated in value, we mean that it will not buy so much; and we do not have in mind any particular commodity; we think of it as less valuable with respect to anything and everything that could be bought. So when money is of high value we mean that it will buy much, or conversely that much of other things must be offered to obtain any given quantity of money. This is evidently nothing else than what people mean by the "purchasing power" of money. We thus arrive at a satisfactory working definition of the value of money and one that agrees closely with the popular concept. *The value of money is the quantity of other things in general that will be exchanged for a given quantity (ordinarily one unit) of it; i.e., its purchasing power.*

How the value of money is measured. But how is the value of money, or its purchasing power, to be measured? The answer must be found in that which expresses at once the relation of money to all other commodities and services. This is price, not the price of any particular thing, but a composite price of things in general, or what is commonly called the "general price level." It is change in the general price level which indicates change in the value of money. A change in a particular price, while other prices remain substantially the same, is presumably due to changes affecting the particular commodity. But when the general price level rises or falls, we may safely assume that the cause is a change in the value of money. High prices and cheap money mean the same thing. Low prices and dear money are synonymous. A change in the general price level means an opposite change in the value of money. Now the general price level — the composite price of things in general — must obviously be based upon some sort of average of the prices of particular commodities and services. There are various kinds of averages and numerous technical problems connected with their calculation, which we shall study later. For the present it is sufficient to know that it is possible to construct "index numbers" which are capable of reflecting with a high degree of accuracy changes in the general price level or the purchasing power of money.

Factors determining the value of money: volume of trade. Our next problem is to discover the forces which determine the value of money and to learn how they operate. The chief service performed by money is the making of exchanges. From the analogy of demand and supply and the value of commodities in general, it should be evident that the value of money is dependent in part upon the volume of exchanges, since it is this that measures the need of money. The more money work there is to be done, as measured by the volume of exchanges, the greater is the need of money, and the greater will be the value of a unit of it; and *vice versa*. As our first step we conclude that the value of money tends to vary directly with the amount of trade.

Quantity of currency. It is by now a commonplace that the value of a unit of anything (*i.e.*, its price) tends to vary inversely with the quantity available for use, the result working out through supply. In a season when there is an unusually large wheat crop we expect, other things being equal, to see the price of wheat comparatively low. A crop failure generally leads to a high price of wheat. So it is with money; the more dollars there are to do a given amount of money work, the less is the value of one dollar; and *vice versa*. In other words the value of money tends to vary inversely with its quantity.

All of the money work however is not performed by money, and in determining the value of money account must be taken of the other forms of currency. The only important one is bank deposits subject to check. The quantity of bank deposits has the same relation to the value of money as the quantity of money itself. The conclusion of the previous paragraph should therefore be modified to read: the value of money tends to vary inversely with the quantity of currency.

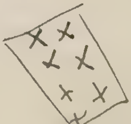
However the quantity of money and the quantity of bank deposits are not wholly independent magnitudes. As we have learned, the amount of deposits which the banks have outstanding is governed in part by the amount of their cash reserves. When the ratio of reserves to deposits is high, the banks are likely to encourage loans and so increase their deposits. When the reserve ratio is low, loans are curtailed and deposits checked. There tends thus to be at least a rough relation between reserves and deposits. And since the division of the total stock of money between bank reserves and other uses depends upon habits of the people which have a certain degree of persistence, we can discern a relation between the total quantity of money and the deposits of the banks. Moreover causes which would lead to increases or decreases in bank deposits, such as business activity or depression, tend also to have the same effect upon the volume of money, principally through changes in the amount of bank notes, and this contributes also to maintaining the relation between money and deposits. Therefore it is not incorrect to say that the value of money tends to vary inversely with its quantity, if only we recognize that the relation is not exact or invariable but may be disturbed by other factors. Its historical verification we shall see later.

Velocity of circulation. We must not forget however that money, unlike most other commodities, is not consumed when used. A coin or note, after facilitating one exchange, is ready with unimpaired efficiency to perform another exchange.[1] The efficiency of a given quantity of money to make exchanges is thus materially affected by its rapidity of turnover; that is, the number of times on the average that a unit of money changes hands during a certain period of time, say a year. It has been estimated that money changes hands on the average in the United States about 30 times a year. This is called the "velocity of circulation" of money. The amount of money work that can be done in a given time is the product of the quantity of money multiplied by its velocity of circulation, and changes in velocity have the same effect on the value of money as proportional changes in its quantity. This principle is of course equally true of other forms of currency, in particular of bank deposits. We conclude therefore that the value of money tends to vary inversely with the velocity of its circulation and the velocity of circulation of bank deposits.

General conclusion. Summarizing our separate conclusions,

---

[1] The fact that paper money wears out is not of significance, since the worn-out notes are always automatically replaced by new ones.



we find that the value of money tends to vary inversely with the quantity of currency and its velocity of circulation and directly with the volume of trade. Since it is the general price level which measures the value of money, varying inversely with it, our conclusion may equally well be stated thus: the general level of prices tends to vary directly with the quantity of currency and its velocity of circulation and inversely with the volume of trade.

The reader may possibly be helped to a clearer understanding of these relations if they are put in the form of an algebraic equation. Let M stand for the average quantity of money in circulation during the year, and let M' denote the average amount of bank deposits subject to check and all other forms of currency. (Practically the other forms are insignificant and may be disregarded.) Let V and V' denote the velocity of circulation per year of money and of other forms of currency respectively. Let T stand for the volume of trade; that is, the total number of units of goods and services exchanged for currency during the year; and let P represent the general price level; that is, the average price paid for all the units of goods and services so exchanged. We then have the following equation:

$$MV + M'V' = PT$$

The quantity of money multiplied by its velocity of circulation is evidently the total amount of money paid in making exchanges during the year. Likewise the quantity of deposits multiplied by their velocity of circulation gives the total amount of payments by transfer of deposits.[1] The sum of these two products is the total amount of payments in all kinds of currency. The volume of trade multiplied by the average price is obviously the total value of all things exchanged for currency. This magnitude and the total amount of currency payments are necessarily the same, and the "equation of exchange," as it is called, is the algebraic expression of this fact.

It is P which measures the general level of prices and (inversely) the value of money, and the effect upon P of any of the other magnitudes is clearly shown by the equation. If we direct our attention to the two most important factors, M and T, it is clear that changes in M tend to cause similar changes in P and changes in T tend to cause opposite changes in P;[1] which is the algebraic expression of the general conclusion that *a change in the quantity of money tends to cause a like change in the general price level and a change in the volume of trade tends to cause an opposite change in the general price level.*

---

[1] Strictly speaking $M'V'$ includes also any other forms of currency that may have been used and which, for simplicity's sake and because of their practical insignificance in the problem, we shall hereafter ignore.

Standard or fiat money. In all of this we have assumed a monetary system based upon standard money, presumably gold. As we have previously seen, there is an automatic adjustment between the quantity of the standard commodity, say gold, which is used as money and that which is devoted to other uses, so that the values of a given quantity of gold coin and the same quantity of gold bullion are the same. The present analysis is therefore quite in harmony with our previous conclusion that the value of any kind of money is the intrinsic value of the standard.

---

[1] *Qualifications.* For the sake of complete accuracy, we have to raise the question whether changes in M, M', and T will always work themselves out on P rather than on each other or on V or V'. A sudden great increase in the quantity of money might have an effect on its velocity of circulation, money passing on the average more slowly because of the greater quantity in people's possession. A sudden great decrease might have the opposite effect, causing the limited quantity of money to be turned over more rapidly. But these effects would be only temporary, and the habits of the people would soon adjust themselves to the new circumstances. As regards the relation between M and M' and T, it is true that a sudden great increase in M or M' would very probably increase temporarily the purchases of the people and so increase T. A corresponding decrease in M or M' might decrease T. But these influences also would be of only temporary potency, the volume of a country's trade depending on other and more fundamental causes.

As regards the volume of trade, the situation is not quite so simple for the reason that an increase in the quantity of goods exchanged is very likely to have a material effect in increasing the amount of money and still more the amount of deposits, while a decrease in trade has the reverse effects. As we have seen, an elastic currency responds to business needs, especially through changes in the amount of bank notes (which are money) and bank deposits. Changes in T thus may cause similar changes in M and M', and such changes will offset the tendency of the changes in T to cause opposite changes in P. However it is certain that in the long run the effects upon M and M' will not be sufficient entirely to counteract the effect upon P, and we come to the conclusion that changes in the quantity of goods exchanged, while part of their effect is upon other quantities in the equation, will nevertheless tend to cause opposite changes in the price level.

These qualifications therefore do not impair the accuracy or usefulness of our conclusion as to the general relation between the quantity of money, the volume of trade, and the value of money.

The case of fiat money is somewhat different in that we have now cut loose from the intrinsic value of any commodity. There is however no qualification to our law of the value of money, which holds of fiat money as well as of a standard monetary system. Fiat money generally develops out of a form of credit money at first redeemable in and as good as the standard money. As its quantity is increased, the monetary system is inflated and, in accordance with Gresham's law, the surplus is drained off in the form of gold, or whatever the standard may be. There may be no depreciation of the credit money in terms of standard money, even after it has ceased to be redeemable and becomes fiat, until the time when its amount exceeds the total original quantity of standard and fiduciary money and when it has driven all the standard money out of circulation. Historically fiat money has always been issued in excess. It has quickly driven standard money out and has become itself the basis of the monetary system. Its value thereafter is determined, as the value of money is always determined, primarily by its quantity and the volume of trade. It always depreciates in terms of the former standard. And the former standard money, say gold, is now a commodity with a price, like any other price, measured in terms of the fiat money unit and affected like other prices as the depreciation of the fiat money raises the general price level. As we have seen, the United States experienced all this during the greenback régime.

Index numbers. The economic law of the value of money or the general price level has been verified time and again by the facts of monetary history. Before passing to this part of our inquiry however we need to acquire a more precise notion of the general price level and some acquaintance with the instrument by which it is measured — the index number.

Of course all prices do not move in unison. Some rise, while others fall, and still others remain stationary. A complete picture of all price changes in any given period would show a kaleidoscopic confusion of thousands of separate and apparently unrelated prices rising or falling or standing still. Yet there is such a thing as a general price movement, just as there is a movement of a whole swarm of bees while the individual bees are dashing about in all directions in the utmost apparent confusion and without seeming relation to each other.

Price relatives and the index number. For a simple example, let us consider the average wholesale prices of three commodities, bituminous coal, wheat, and copper, in the three years 1913, 1914, and 1915. Here are the facts:

PRICES

	1913	*1914*	*1915*
Coal	$1.27 per ton	$1.17 per ton	$1.04 per ton
Wheat	.91 per bu.	1.04 per bu.	1.34 per bu.
Copper	.15 per lb.	.13 per lb.	.17 per lb.

Evidently the price of coal declined during this period, the price of wheat rose, and the price of copper fell from 1913 to 1914 and rose in the next year. To measure these separate price movements we may call one year, say 1913, the base year and express each price of each commodity as a percentage of the price of that commodity in the base year. For example the prices of coal were: in 1913, $1.27; in 1914, $1.17; in 1915, $1.04. Calculating the percentages we have: $\frac{1.27}{1.27} = 1$, or 100%; $\frac{1.17}{1.27} = .92$, or 92%; $\frac{1.04}{1.27} = .82$, or 82%. If then we let 100 stand for the price of coal in 1913, the price in 1914 relative to it will be 92, and the 1915 price, 82. Such numbers are called *price relatives*. By similar calculations we find the price relatives for wheat and copper with respect to their prices in 1913, and the whole result is shown in the following table:

PRICE RELATIVES

	1913	*1914*	*1915*
Coal	100	92	82
Wheat	100	114	147
Copper	100	87	113

Such a table might be continued to include similar price relatives for all other important commodities, giving thus a record of the price changes of each separate commodity. But what we now desire is to know how prices in general have changed; we are seeking a record of changes in the general price level. For this we need a number which shall express for each year the general or average price relative, evidently the average of the separate price relatives for each year. Such a number is called an *index number* and may be defined thus: *An index number is a number which expresses the general price level for any given year relative to the general price level of some year which is taken as the base year and whose general price level is represented by 100.*

The simple arithmetic mean. Now there are many kinds of averages, and the particular average chosen will determine the character of the index number. For example, we might employ the simple arithmetic mean of the separate price relatives of each year. In our example the average for 1913 is of course 100. For 1914 the simple arithmetic average is $\frac{92 + 114 + 87}{3}$, or 98. For 1915, it is 114. Index numbers have been constructed on this principle, but they have a serious defect. The several commodities which enter into trade are not of equal importance, and an index number which takes no account of such differences is not an accurate record of the changes in the general price level or the purchasing power of money. The value of money is a good deal more affected, for instance, by a change in the price of wheat than by an equal change in the price of clover seed.

The weighted average. To avoid this defect of the simple arithmetic mean it is necessary to use a weighted average which takes proper account of the relative importance of the several commodities. For example suppose that in a given year the price of corn increased four per cent, while the price of butter increased ten per cent, and that corn was twice as important as butter. The simple arithmetic average would show a combined increase of seven per cent; $\frac{4 + 10}{2} = 7$. But this is not a true picture, since no account is taken of the greater importance of corn. The average is corrected by putting the figure for corn in twice; thus $\frac{4 + 4 + 10}{3} = 6$. This is a weighted average of the percentage increases. Weighting may be accomplished in various ways, the most satisfactory method being to weight the price movement of each commodity according to the total value of that commodity exchanged in a given year. Thus if twice as much money is expended for commodity A as for commodity B in the selected year, this is taken to indicate that A is twice as important as B and should be given double the weight of B in the index number. The simplest way to accomplish this is to calculate first the total value of all the goods exchanged in a given year and then to compare with this the value of exactly the same quantities of these goods reckoned at the prices of each of the other years. To make the method clear, let us return to our previous example. Here are the facts for the first year:

YEAR 1913

	Prices in 1913	*Quantities exchanged in 1913*	*Values at 1913 prices*
Coal	$1.27	447 million tons	568 million dollars
Wheat	.91	555 million bu.	505 million dollars
Copper	.15	812 million lbs.	122 million dollars
Total value			1,195 million dollars

It cost $1,195,000,000 to buy these quantities of these commodities in 1913. If the same quantities of these commodities cost more or less in 1914, it must have been because of changes in their prices, and if we could determine what the cost was in 1914 we should have a very good measure of the average movement of these prices. This may be readily calculated, as follows:

YEAR 1914

	Prices in 1914	*Quantities exchanged in 1913*	*Values at 1914 prices*
Coal	$1.17	447 million tons	523 million dollars
Wheat	1.04	555 million bu.	577 million dollars
Copper	.13	812 million lbs.	106 million dollars
Total value			1,206 million dollars

Evidently commodities which cost $1,195,000,000 in 1913 were worth $1,206,000,000 in 1914. There must have been a small rise in the average price level of these three commodities.

For the year 1915 we obtain similar results, as follows:

YEAR 1915

	Prices in 1915	*Quantities exchanged in 1913*	*Values at 1915 prices*
Coal	$1.04	447 million tons	465 million dollars
Wheat	1.34	555 million bu.	744 million dollars
Copper	.17	812 million lbs.	138 million dollars
Total value			1,347 million dollars

Again there has evidently been a rise in the average price level.

An index number of the aggregative type. The final step in deriving a real index number of the prices of these three commodities is to express the total value for each year as a percentage of the total value in 1913. Thus: $\frac{1,195}{1,195} = 1.00$, or 100%; $\frac{1,206}{1,195} = 1.01$, or 101%; $\frac{1,347}{1,195} = 1.13$, or 113%. These percentages are the index numbers and the series may be thus expressed:

Year	*Index number of prices*
1913	100
1914	101
1915	113

This represents an index number of the *aggregative* type. The year 1913 is the base year, its index number being 100. The weighting is according to the values exchanged in 1913. Though the same in our example, the year chosen for the weighting need not necessarily be the base year.

Practical index numbers. The example used in the foregoing explanation was purposely made simple by employing only three commodities. A practical index number must take account of a large number of articles and may be calculated for any number of years, but no new principle is thereby involved. It must also be noted that the weighted arithmetic mean, which we have used, is by no means the only possible average. The geometric mean has certain advantages, and the theoretically ideal formula is more complicated than either of these. It is not necessary for us to follow this path further. Fairly satisfactory results are obtained from each of several formulas. The aggregative type of index number which we have studied serves well to bring out the essential principles of index numbers in general, and it is also one of the most important and on the whole probably the best of the types in actual practical use.

One of the best index numbers and the one most generally used in the United States is that published by the United States Bureau of Labor Statistics. This index number was first published in 1902, since which time it has undergone various changes from an unweighted mean of price relatives to its present form, which is an example of the aggregative type described above. It shows the index number of wholesale prices for each year from 1890 to the present and for each month since January, 1913. The number of commodities used has varied from time to time. In the first year, 1890, there were 192 commodities or separate series of quotations; in 1914 the number had increased to 297; in 1925 it was 404. The base year is now 1913, and the weighting, utilizing the latest Census reports of quantities exchanged, is at present made in ac-

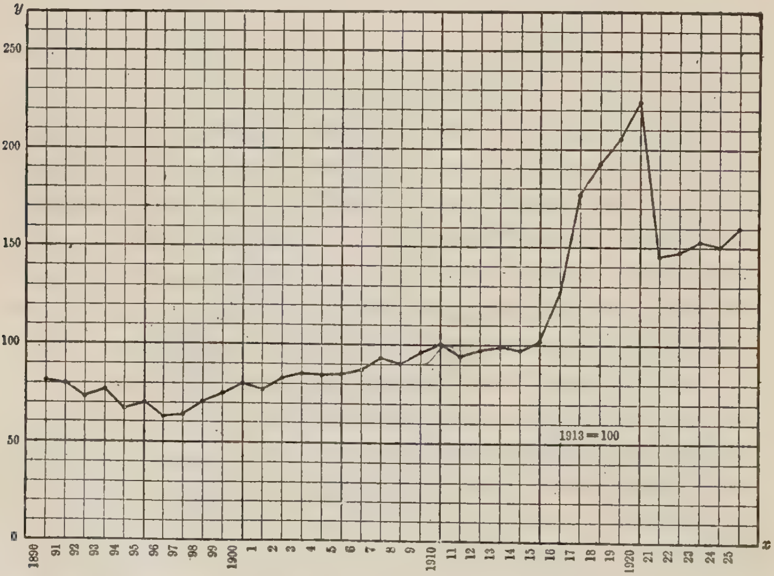


FIG. 43. — Index Number of Wholesale Prices in the United States, 1890–1925.

cordance with the values exchanged in the year 1919. The table below gives the annual series from 1890 to 1925, to which we have added the reciprocals which show the purchasing power or value of the dollar (calling its value in 1913, 100).

Wholesale Prices and Value of Money in the United States
1890–1925

Year	*Index Number of Prices*	*Purchasing Power of the Dollar*
1890	81	123
1891	80	125
1892	75	133
1893	77	130
1894	69	145
1895	70	143
1896	67	149
1897	67	149
1898	70	143
1899	75	133
1900	81	123
1901	79	127
1902	84	119
1903	86	116
1904	86	116
1905	86	116
1906	89	112
1907	94	106
1908	90	111
1909	97	103
1910	101	99
1911	93	108
1912	99	101
1913	*100*	*100*
1914	98	102
1915	101	99
1916	127	79
1917	177	56
1918	194	52
1919	206	49
1920	226	44
1921	147	68
1922	149	67
1923	154	64
1924	150	67
1925	159*	63*

* Average of monthly index numbers of January–September.

As an example of the monthly series of index numbers of the United States Bureau of Labor Statistics, the following figures are given for 1924 and 1925, to which we have again added the reciprocals showing the monthly changes in the purchasing power of the dollar.

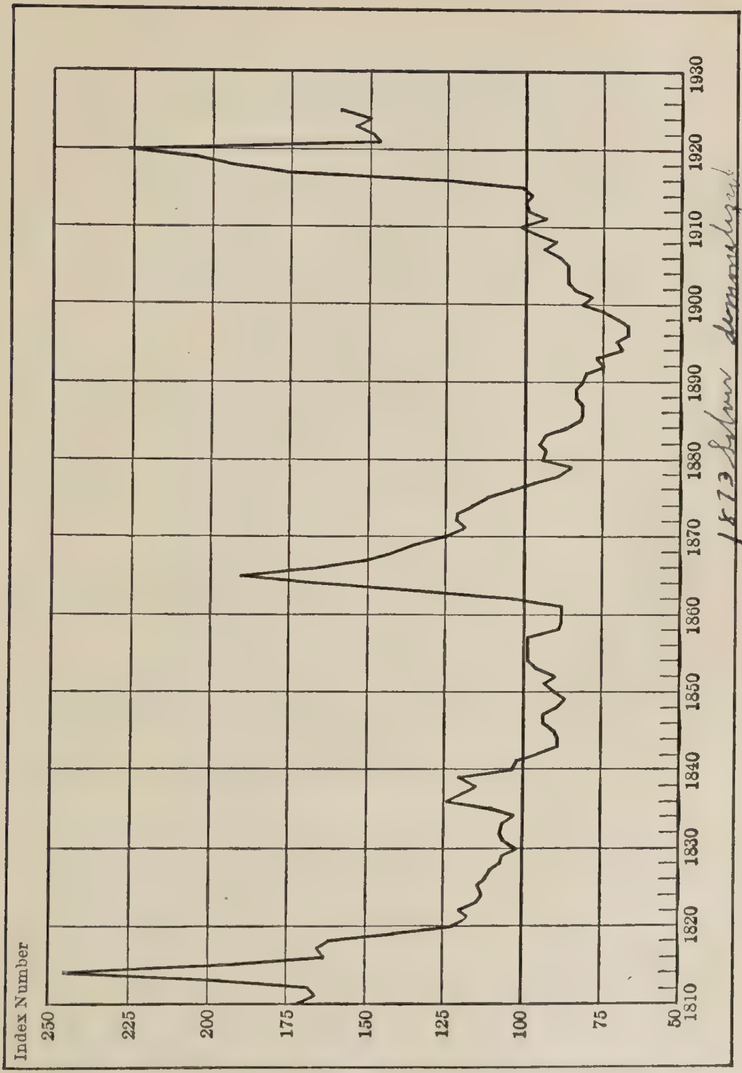
WHOLESALE PRICES AND VALUE OF MONEY IN THE UNITED STATES, BY MONTHS, 1924–1925

	1924		*1925*	
	Index Number of Prices	*Purchasing Power of the Dollar*	*Index Number of Prices*	*Purchasing Power of the Dollar*
January	151.2	66.1	160.0	62.5
February	151.7	65.9	160.6	62.2
March	149.9	66.6	161.0	62.1
April	148.4	67.3	156.2	64.0
May	146.9	68.0	155.2	64.4
June	144.6	69.1	157.4	63.5
July	147.0	68.0	159.9	62.5
August	149.7	66.8	160.4	62.3
September	148.8	67.2	159.7	62.6
October	151.9	65.8		
November	152.7	65.4		
December	157.0	63.6		

For many purposes the most useful way to present an index number is by means of a graph or curve. The Bureau of Labor Statistics' annual index number of wholesale prices is thus shown in Figure 43, on page 488, the point for 1925 being based upon the average of the monthly index numbers of January–September.

There are many other practical index numbers, computed and published by various agencies in America and Europe, using various formulas and relating to the prices of selected groups of commodities in the respective countries. The United States Bureau of Labor Statistics publishes periodically index numbers for the leading foreign countries. An interesting weekly index number of American wholesale prices is computed and published in the newspapers by Professor Irving Fisher.

Historical verification. We are now in a position to appreciate the historical data illustrating and confirming the economic law of the general price level and the value or purchasing power of money. These data are at hand in embarrassing abundance. Many of the incidents referred to in previous chapters on money will now be recognized as illustrations of this law. It will suffice for our present purpose to cite only a few examples, taken from the monetary history of the United States since 1860. As an aid to



1873 Silver demonetized!

FIG. 44. — Index Number of Wholesale Prices in the United States, 1810–1925.[1]

---

[1] The index numbers for this diagram were supplied in part by The National Industrial Conference Board. Index numbers from 1810–1840 were compiled by the Russell Sage Foundation; from 1840–1925 by the United States Bureau of Labor Statistics.

this inquiry we have combined several series of index numbers to give the graph in Figure 44 (on the preceding page).

An increase in the quantity of money: fiat money. At the end of the fiscal year 1861 (that is, June 30, 1861) the total monetary stock of the United States, as calculated by the Treasury statisticians, was 488 million dollars.[1] During the next three years there were issued, in connection with the financing of the Civil War, 447 million dollars of United States notes or greenbacks, 23 millions of similar notes of denominations less than one dollar, called "fractional currency," and 169 millions of various other forms of government fiduciary money. These issues, in combination with certain other monetary changes, brought the total monetary stock to 1,063 million dollars in 1864. Thus in three years the quantity of money was more than doubled. The result upon prices and the value of money was in harmony with the economic law. The wholesale price level nearly doubled, rising from 88 in 1861 to 167 in 1864. Reciprocally the purchasing power of the dollar declined from 114 to 60.

The greenbacks were from the first irredeemable fiat money. Except on the Pacific coast, they expelled most of the gold and silver from circulation and were themselves the basis of the monetary system. They depreciated in terms of gold, the depreciation being generally fairly proportional to the rise in the general price level. For example, it required during 1864 on the average 203 dollars in greenbacks to buy 100 dollars in gold coin or an equivalent weight of gold bullion.

A decrease in the quantity of money. During the next four years (1864–1868) conditions were reversed. There were material reductions in the volume of greenbacks, state bank notes, other forms of fiduciary money, and gold coin, offset in part by increases in national bank notes and fractional currency, the net result being a decrease in the total monetary stock from 1,063 million dollars to 888 millions. The decline in the quantity of money tended to lower the general price level, an effect which was also promoted by the growth in trade after the war. The joint effect was a drop in the index number from 167 in 1864 to 142 in 1868. Considered as a change in the value of the dollar, this indicates an advance in its purchasing power from 60 to 70. The depreciation of the paper money in terms of gold also became less. In 1868 it took 140 dollars in greenbacks to buy 100 dollars in gold, whereas it required 203 dollars in 1864.

---

[1] In the present discussion the figures of total monetary stock of the United States do not include the "minor coins," though they do include the token coins of silver. For a more refined investigation it would be desirable to use, not the total monetary stock, but only such money as is in "circulation" or is the basis of other currency in circulation. However for the sake of the broad conclusions which we are seeking and in view of the fact that we are taking account of money only, not deposits, the use of figures of the total monetary stock is sufficiently accurate.

An increase of trade. The effect of increased trade stands out even more clearly in the history of the next decade. From 1868 to 1879 the total quantity of money remained fairly stable as compared with the preceding years. There was a gradual increase, from 888 to 1,034 millions. This increase in the quantity of money, other things being equal, would presumably have caused a moderate rise in prices. But other things were not equal. The country was growing rapidly in population and wealth, and trade was expanding. An increase in the volume of trade, as the reader will recall, tends to cause a fall in prices, and the increase during this period was so great as to outbalance the opposite tendency of the increase of money. Hence the net result was a fall of prices. From 1868 to 1879 the index number of wholesale prices fell from 142 to 85. Reciprocally the purchasing power of the dollar rose from 70 to 118.

The end of fiat money. During this period the greenbacks continued their rise toward par in terms of gold, which was finally reached in the closing days of 1878, and on January 1,1879, the government began redeeming them in gold on demand. Thus ended the fiat money régime which had continued for seventeen years; the metallic standard was restored. By this time the volume of greenbacks had been permanently fixed at 347 millions, and the fractional currency, state bank notes, and miscellaneous forms of fiduciary currency had been eliminated from the monetary system, which for the following generation comprised only coin, greenbacks, and national bank notes.

Increases in money and trade. The next chapter in our history takes us to about the year 1897. During these eighteen years the monetary stock nearly doubled, being 1,907 million dollars at the close of 1897. The increase was due to the silver dollars of the Bland-Allison and Sherman acts and to a material gain in gold, overcoming a decline of nearly 100 millions in national bank notes. This gain in the quantity of money was however, after the first few years, not enough to counteract the phenomenal expansion of the country's trade which, gathering momentum after the Civil War, has progressed to the present day. These two influences working against each other in the equation of exchange, the net result was a quick rise in prices in 1880 and thereafter a steady decline, which was not permanently checked till after 1897. The index number of prices in this year was 67, which marks the low record to the present day. The purchasing power of the dollar in 1897, on the basis of 1913, was 149.

Monetary inflation through gold. The period from 1897 to the present has been on the whole one of monetary inflation. Up to the beginning of the World War in 1914 the inflation, strangely enough, was not due to paper money issues but to gold, in marked contrast with previous and subsequent periods of monetary expansion. True the circulation of the national banks increased from 231 millions in 1897 to 759 millions in 1913, but this was a moderate increase alongside the extraordinary growth of our gold currency from 696 to 1,871 millions. This was the result of a tremendous gain in the output of the world's gold mines, a phenomenon of the utmost importance in view of the well-nigh universal adoption of the gold standard. We can well afford to give some attention to the history of the production of the world's chief monetary metal.

The world's production of gold. It is estimated that the average annual production of gold throughout the world four centuries ago (from 1493 to 1520) was a little less than 4 million dollars worth, and that during the next three and a quarter centuries the rate of production increased gradually to about 13 millions in the decade 1831–1840. Then came an unexampled increase, due principally to the rich discoveries in California and other western states, which in twenty years had multiplied the world's annual production tenfold; it averaged 134 millions in the five years, 1856–1860. After this climax production fell off gradually to a little less than 100 millions in 1881–1885. Then began a new development, resulting from fresh discoveries of gold in South Africa, Canada, and Alaska, and to improved processes of extracting the gold from low grade ores. Thus in a few years the world's annual production reached a new high level, which has been substantially maintained ever since. In the first year of the twentieth century, 261 million dollars worth of gold were added to the world's stock. In 1906 the year's production exceeded 400 millions, and it stood well above that mark for the next ten years and has been only slightly less in the years since 1917. For the past twenty years therefore the world's store of gold has been accumulating at thirty times the rate at which it was being built up during the eighteenth and first half of the nineteenth century. The gold produced in each of the years from 1908 to 1916 exceeded the total production of the first forty years of the nineteenth century.[1]

---

[1] These facts are shown in more detail in the following table:

WORLD'S GOLD PRODUCTION

	Annual Average for the Period	
Period	*Fine ounces*	*Value*
1831–1840	652,291	13,484,000
1841–1850	1,760,502	36,393,000
1851–1855	6,410,324	132,513,000
1856–1860	6,486,262	134,083,000
1861–1865	5,949,582	122,989,000
1866–1870	6,270,086	129,614,000
1871–1875	5,591,014	115,577,000
1876–1880	5,543,110	114,586,000
1881–1885	4,794,755	99,116,000
1886–1890	5,461,282	112,895,000
1891–1895	7,882,565	162,947,000
1896–1900	12,446,939	257,301,100
1901–1905	15,606,730	322,619,800
1906–1910	20,971,575	433,520,960
1911–1915	22,213,810	459,200,175
1916–1920	18,966,758	392,074,674
1921	15,974,962	320,231,792
1922	15,451,945	319,420,063
1923	17,750,765	366,940,884

Effect on the price level. This deluge of gold into the world reservoir had its inevitable effect on the monetary system of the United States. In the sixteen years from 1897 to 1913, the total stock of money doubled again (rising from 1,907 millions to 3,720 millions), the three principal items in the increase being: subsidiary silver coins, 99 millions; national bank notes, 528 millions; and gold, 1,175 millions. This addition to the monetary stock was more than enough to counteract the continuing growth in the country's trade, and the result was a turn in the movement of prices. The general price level, which at the end of a long recession had reached its low point in 1897 (index number, 67), now began a rise which carried it, almost without interruption, to 100 in 1913 (the base year of the series). This meant a fall in the purchasing power of the dollar from 149 in 1897 to 100 in 1913.

Monetary inflation of the World War. The World War ushered in a period of inflation exceeding anything in the previous history of the United States. The steady outpouring of the world's gold mines continued, but now the United States began to receive more than her normal share.[1] For eleven years (1913–1924) the gold coin and bullion in the United States monetary system grew at the average rate of 238 million dollars a year. Whereas at the end of 1913 the United States held a little less than two billion dollars of gold out of a total world stock of about eight and a

---

[1] The following table, based principally upon the reports of the Director of the United States Mint, shows the result of the World War upon the gold money of some of the important nations:

Gold money held by	*December 31, 1913*	*About Dec. 31, 1923*
Russia	$1,011,500,000	$ 45,043,000
Germany	915,700,000	119,300,000
Canada	142,500,000	227,292,000
Italy	265,000,000	215,697,000
Holland	60,900,000	233,876,000
Argentina	292,600,000	472,161,000
Spain	92,500,000	487,687,000
Japan	64,963,000	602,188,000
France	1,200,000,000	709,479,000
Great Britain	830,100,000	759,174,000
U.S.A.	1,904,700,000	4,247,201,000
World Total	$8,240,000,000	$8,925,000,000

quarter billions (even then far exceeding any other nation), it has been estimated that by the close of the year 1924, the United States held the huge total of 4,491 millions, having more than doubled her holdings and having gained possession of nearly half the total gold stock of the world. Our gold stock alone in 1924 was considerably more than the total of all kinds of money ten years earlier. Added to this was a great increase of federal reserve notes, which, appearing first in 1914, reached a maximum of 3,406 millions in 1920 and then declined to 2,339 millions in 1924. The total monetary stock thus rose from 3,720 millions in 1913 to 8,746 millions in 1924.

Soaring prices the result. If the reader will turn back to the series of annual index numbers on page 489 he will find the record of soaring prices which we should expect as a result of such an increase in the quantity of money. By 1918 the price level had doubled; it reached a maximum of 226 for the year 1920. The highest monthly index number recorded was 247 for May, 1920. Thereafter it declined quite suddenly to 147 for the year 1921, remaining not far from 150 since then; this new level being just about fifty per cent higher than prevailed during the five or six years immediately preceding the World War. The purchasing power of the dollar in 1920 was only 44 (as compared with 100 in 1913 and 102 in 1914) and since 1920 it has stood at about 67.

Lower prices since 1920. The sudden drop of prices in 1921 and the lower level which has since continued do not appear to be matched by corresponding changes in the quantity of money. Without entering into all the complicated factors of the problem, it should be noted that the increase in money since 1920 (to 1924) has all been due to gold; in fact there was a decrease of nearly a billion dollars in the other forms of money. Now the influx of gold has gone almost entirely into the possession of the federal reserve banks and the United States Treasury, where, in accordance with a conservative policy, it has to a considerable extent been held as an idle store and not permitted to have the effect upon the money in circulation which might otherwise have occurred.

Conclusion. This line of inquiry might well be pushed on into an investigation of the fluctuations in bank deposits and their relation to the price level. This the limits of our space will not permit, but the reader has now before him enough of the facts of our monetary history fully to verify the economic law of the relation between quantity of money and volume of trade and the value of money and to convince him that this law is no mere abstraction but a living truth of the utmost practical significance to mankind.

Far-reaching effects of monetary fluctuations. As to this latter conclusion some further observations will not be out of place. Fluctuations in the value of money are fraught with the most serious consequences to all the people, inflicting fortuitous loss upon some and bestowing undeserved gain upon others. Rights to future payments are normally expressed in terms of the country's monetary unit. When the time for payment comes the debtor will pay and the creditor will receive money whose value may be quite different from that which existed when the agreement or contract was made. A farmer borrowed $10,000 at a time when wheat was worth $1.00 a bushel. If before the time for payment arrives there has been such a fall in the value of money that prices, including the price of wheat, have generally doubled, the farmer will be able to pay his debt in dollars worth only half as much as those he borrowed. When he made the loan, it required the sale of 10,000 bushels of wheat to produce $10,000. Now he need sell only 5,000 bushels to obtain the $10,000 needed to pay his debt. The debt has been virtually halved, to the advantage of the farmer and of course to the corresponding loss of the one from whom he borrowed. On the other hand if, in consequence of a doubling of the value of money, the price of wheat had fallen to fifty cents the result upon both parties would have been the reverse. This illustrates the general principle of the effect of fluctuations in the value of money upon debtors and creditors.

All investments which consist of rights to receive stated sums of money, such as bonds, mortgages, annuities, insurance, etc., are subject to the same principle. The investor is compelled in spite of himself to gamble on the value of the dollar. If it goes up, the dollars he receives will buy more; if it falls, his dollars will buy less. Such investors were fortunate gainers from this cause during the long period of falling prices that ended in 1897. Since then the most conservative investors have seen their resources dwindle from year to year as the index number of prices rose. The man who in 1904 put money in the savings bank, where he left it to accumulate at five per cent till 1918, suffered a loss through the shrinkage of the dollar sufficient to wipe out all the gain from his compound interest for the fourteen years. These losses on investments are not due to any default on the part of the debtor. The same loss would result if the actual money were held intact in a safe.

Those whose incomes are in the form of salaries and wages are affected by fluctuations in the value of money the same as are creditors. They have agreed to render services in return for a stated number of dollars, which as they are received may be worth more or less than when the agreements were made. Only as agreements expire and new ones are made can the parties undertake to make new terms conforming to the new monetary value, and these new agreements are equally subject to the vicissitudes of future changes.

It is thus that adjustment lags behind the fluctuations of the monetary standard on all agreements except those which are for immediate payment or for relatively short terms. Here there is the current adjustment, through raising and lowering prices, which produces those fluctuations in the general price level by which the community adjusts itself as best it may to the wabbling monetary standard.

Collapse of fiat money. Under a fiat money régime there is no limit to the possible decline in the value of money. The American continental bills of credit became practically worthless and were finally redeemed, such as had not been lost, destroyed, or thrown away, at one cent on the dollar. The original currency of Soviet Russia lost all value. The German currency was inflated during and following the World War until its depreciation was impossible to reckon and finally received official recognition at the grotesque ratio of one trillion to one. This meant the practical wiping out of all debts and the destruction of all investments and savings which were in the form of rights to receive definite sums of money. It impoverished millions of the people and enriched others. It is difficult to imagine such colossal wholesale robbery. If all the burglars and highwaymen in the country were given free rein for a decade their best efforts could not remotely approach this accomplishment.

The gold standard. The reader will not make the mistake of supposing that such iniquities come only from irredeemable paper money. Our study has fully demonstrated the fluctuating character of any monetary system, no matter how firmly established on the foundation of gold. All that can be said is that the gold standard gives some stability and prevents the extremes of depreciation. Gold is itself a commodity, useful for other purposes than to serve as money. It is produced only by incurring costs which must not for long be greater than its value. Its quantity cannot be expanded indefinitely, and it cannot lose all value. On the whole it has proved itself the most satisfactory of any of the commodities that might be chosen as the monetary standard. And yet in three generations in the United States we have seen a gain of forty per cent and a loss of seventy per cent in the value of the gold dollar itself.

Is a more stable standard possible? As a general rule those who have understood the economic principles of money and have sought the maintenance of the most stable standard possible have had their hands fully occupied in repelling the attacks of those who sought to displace the gold standard by something far less stable. Only recently has there been any considerable recognition of the imperfections of the gold standard or opportunity to seek something that should be better. At present the number of those who realize the situation is growing rapidly, and there is considerable sentiment in favor of a new monetary standard which shall be so far as possible free from fluctuation in terms of the general price level. Such a standard must of course be based in some way upon an index number of prices. Perhaps the leading proposition today is the so-called "stabilized dollar." Without going into technical details, we may note that this plan would make the monetary unit — the dollar — not a certain weight of gold as now, but a variable weight of gold, so adjusted from time to time as to keep its purchasing power virtually constant. Gold coin would not be used in actual circulation; in fact very little is so used now. The government would be as now the custodian of the country's monetary gold. There would be in circulation gold certificates and other forms of money reading in terms of dollars (not necessarily much different from the present forms), all of which would be redeemed in gold by the government on demand. The amount of gold to be given for a dollar would be officially determined from time to time on the basis of the index number of prices, so that a gold dollar would always purchase about the same quantity of goods in general. Variations of individual prices with respect to each other, like the movements of the separate bees in the swarm, would in no wise be affected. But the movements of all prices in general would be stopped; the swarm as a whole would become stationary. This plan, especially if simultaneous adoption by most of the leading nations be assumed, is undoubtedly based on sound economic principles. Whether political obstacles would prove unsurmountable is a serious question, and in any event its practical adoption must await further public education.

CHAPTER XXV

THE BUSINESS CYCLE

Even the most casual observer cannot have failed to notice that business as a whole has its " ups and downs," its periods of prosperity and depression — times when most business men are making profits and are optimistic as to the future and other times when even the strongest firms find it necessary to sail close to the wind in order to avoid the disaster that has already engulfed the weakest. Careful observation seems to show that there is a fairly regular sequence of events in business; prosperity is followed by crisis, the crisis gives way to depression, more or less prolonged, this in turn is followed by a period of recovery, which paves the way for a renewal of intense activity. It is this oscillating aspect of business activity which has given rise to the term " business cycle."

The price level and the business cycle. The outward manifestation of the events included within the business cycle is a complete revolution in price movements; rising prices are followed by falling prices, recovery in prices, and then, at the beginning of the new cycle, rapidly rising prices once more. This cyclical movement of prices should be differentiated from other price movements; namely, those arising from sudden inflation or deflation of the currency and the slower changes in the price level which take place over long periods of time.

In the diagram below, repeated from the preceding chapter, are given the index numbers of wholesale commodity prices in the United States from 1810 to the present time. We note in the first place that there are three periods in which the price level rose with extraordinary rapidity — in the second decade of the nineteenth century, in the Civil War period, and during the World War. In each case the cause was a marked inflation of the currency. In the second place we can see well defined trends in

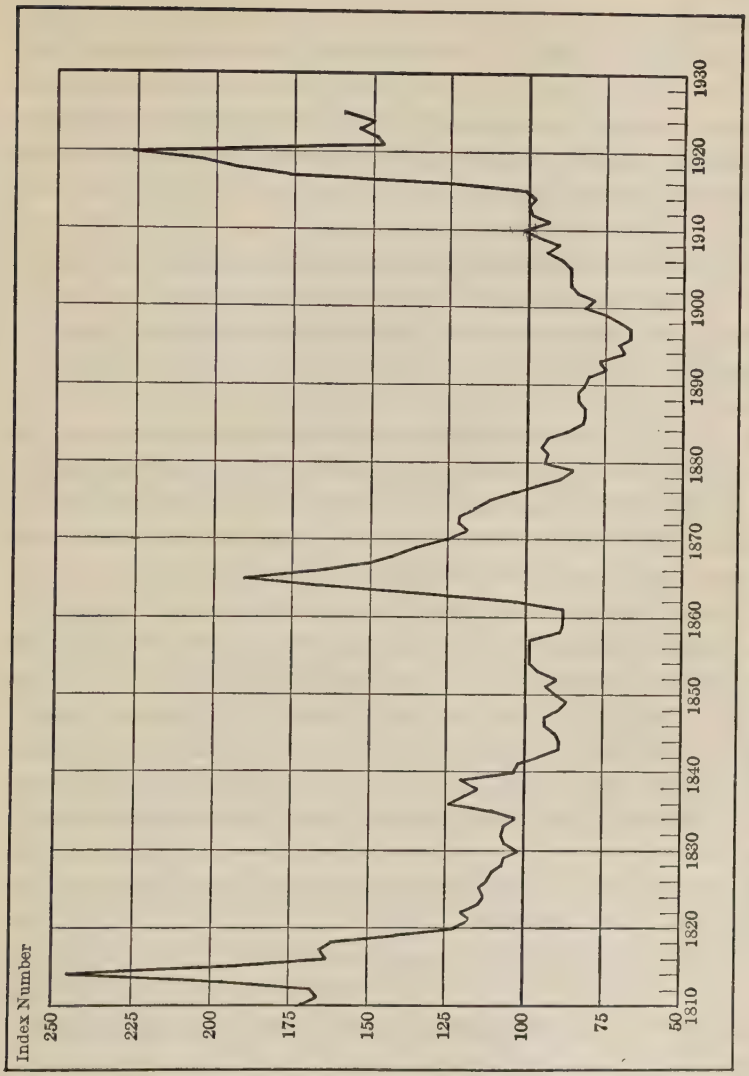


FIG. 45. — Index Number of Wholesale Prices in the United States, 1810–1925.[1]

prices over fairly long periods of time. From the high point reached in the second decade the general trend of prices was downward until the upward movement started in the early forties; from the high point in the sixties the general trend of prices was downward until 1896, and from 1897 on the trend was upward until the peak in 1920. To such a general movement in one direction over a broad sweep of time the term *secular trend* is given. The secular trend is the result of all the forces, direct and indirect, which tend to cause changes in the quantity of currency and the volume of trade and so determine the general price level over long periods.[1]

---

[1] From data supplied in part by the National Industrial Conference Board.

It will however be seen that the downward or upward movement of prices was not unbroken in any one of the periods; at frequent intervals the general trend was interrupted by a shorter upward or downward swing in prices. These variations in prices which take place within the broader price movement are the evidence of the business cycle. But the business cycle is much more than a revolution in prices; every aspect of business — production, profits, wages, investment, interest, etc. — is affected in one way or another during the course of the cycle. We are led therefore to inquire into the sequence of events in the typical cycle.

Phases of the business cycle. Every business cycle exhibits at least two well-defined phases, the period of prosperity and the period of crisis or liquidation, and most business cycles have two other phases which follow the crisis, a period of depression and a period of recovery. Assuming the four-phase cycle to be the normal case, we may begin our description with the period of prosperity, reminding the reader that this stage starts with a heritage from the stage immediately preceding it and that it cannot therefore be considered as fully portrayed until we have completed our description of all phases of the cycle.

The period of prosperity. As a result of the slowing down of business activity in the preceding periods, the stocks of goods of all kinds have become depleted. At the beginning of the period of prosperity, as consumers make their wants known, there is a rise in the prices of some goods, spreading gradually and affecting a constantly greater number of commodities. This increase in demand is passed on by the middlemen to the producers of these goods and by them to the manufacturers of producers' equipment and to the producers of raw materials. The activity thus engendered among producers in general leads to greater employment of labor and in the long run to higher money wages. This means of course greater purchasing power in the hands of consumers and a further stimulation to the producers of both consumers' and producers' goods, and the sporadic rise in prices becomes a general rise. Prosperity once started seems to breed prosperity; it becomes cumulative in its effects and creates a spirit of optimism which exerts a powerful influence on business activity.

---

[1] This subject has been fully considered in Chapter XXIV.

In large measure this optimistic spirit rests on a solid foundation of enlarged profits. Middlemen can make profits by the simple process of buying goods and awaiting the rise in price, and manufacturers also have unusual opportunities for profit making. Not only is the volume of production larger — and therefore the presumption of an increased volume of profits — but the margin of profit tends to increase. Prices do not rise together; some of them are stationary, fixed by contract, such as rentals and the interest rates on bonds already issued, while others are sluggish, moving upward less rapidly than other prices. Wages and interest rates on short-time loans are the most conspicuous examples of these. On the other hand some of the costs incurred by a manufacturer may increase more rapidly than the price of his product, as for example the cost of raw materials and of equipment. Yet on the whole the prices of the goods and services which the typical manufacturer buys rise less rapidly than the prices of those which he sells, and there thus remains a wider margin of profit, which invests the future with a roseate glow and leads the manufacturer to speed up production and expand his productive facilities.

Bank credit. A part, perhaps the major part, of the profits from industry may be reinvested in the business in the form of buildings, equipment, and so on, but in order to secure the desired expansion of facilities most entrepreneurs have to look outside for funds. To a large extent these are secured by the issue of new securities, but business men must have recourse also to the commercial banks for assistance. The banks therefore stand in a pivotal position, for not only do they make loans to the active managers of enterprises but in many cases they finance their clients in their purchases of corporate securities. The importance of the operations of the banks lies in the fact that by extending credit they place purchasing power in the hands of enterprisers, making possible an increased demand for raw materials, equipment, and labor. Continuously rising prices, which are an essential feature of the period of prosperity, can be maintained only if this demand continues to grow, and this can take place only through an increase in the available quantum of purchasing power. Commercial banks by creating bank deposits furnish the necessary purchasing power, and we therefore conclude that the extension of credit by banks is a necessary feature of the period of prosperity. Credit expansion, while it does not initiate the period, makes its continuation possible.

Prices of securities. Among other features of this period we should note that the prices of stocks, particularly of industrial stocks (as distinguished from public utilities, for example), tend to soar. Aside from the normal rise due to the revaluation of shares on the basis of enlarged earnings, the optimistic speculation of a " bull " market forces prices up.

Checks to prosperity. Sooner or later however checks to the tide of prosperity appear. Wages rise as competition among employers for labor becomes keen, and labor becomes less efficient as the need of retaining a particular job becomes less important. Labor troubles become a factor to be reckoned with, for in times of greatest prosperity labor is apt to be most highly organized, to have the most ample resources, and to be fully aware of the disinclination of an employer to allow a strike to close his plant. The increase in the volume of bank loans finally forces the banks to curtail the volume of new loans and renewals, and to do this they raise the discount rate and become more exacting in their loan policy. In the diagram below the interest rates on 60 to 90 day time loans in New York City are given for the years 1919–1921. The break in commodity prices which marked the beginning of our most recent crisis came in May, 1920, and the year 1919

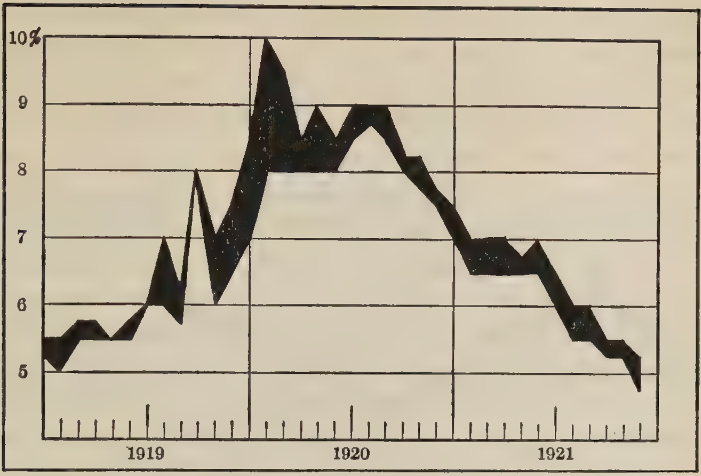


FIG. 46. — Interest Rates on 60 to 90 Day Time Loans in New York City. (Rates shown are the range, high and low, for the first full week in each month, week ending Saturday.[1])

was therefore the last year of the period of prosperity. It will be noted that, starting at fairly moderate rates in the first month of 1919, the rates rose sharply in the last months and attained their greatest height in February of 1920.

Furthermore costs which have been fixed by contract may rise also. Leases expire and can be renewed only on less favorable terms. In the course of time the market for securities becomes congested and new issues of bonds, put out either for the purpose of refunding old issues or to secure funds for additional equipment, cannot be sold except at a relatively high rate of yield.

Costs of production therefore are rising, and they rise more rapidly than the price of the product which the manufacturer is making. Hence the margin of profit tends to disappear, and as costs catch up with prices the period of prosperity nears its close.

The behavior of the market for securities usually gives an indication of the impending change in the tide of business affairs. As profits diminish, stock quotations, though still high, show a tendency to waver. The market becomes somewhat erratic, with bursts of speculative activity breaking the general dullness, and as the period draws near its end there is likely to be a more or less general fall in the prices of stocks. This is shown in the diagram below. The trend of the prices of industrial stocks in 1919 was upward until the last quarter of the year, when prices broke and continued falling until about the middle of 1921. It will be noted that the break in the prices of industrial stocks anticipated by several months the break in general commodity prices. Inasmuch as the prices of stocks reflect the opinion of speculators and investors as to the probable profits of the future, this is to be expected, as is also the opposite situation in the period of recovery, when the rise in the prices of stocks generally anticipates the period of prosperity.

---

[1] Derived from tables given in *Statistical Abstract of the United States*, 1923, p. 664.

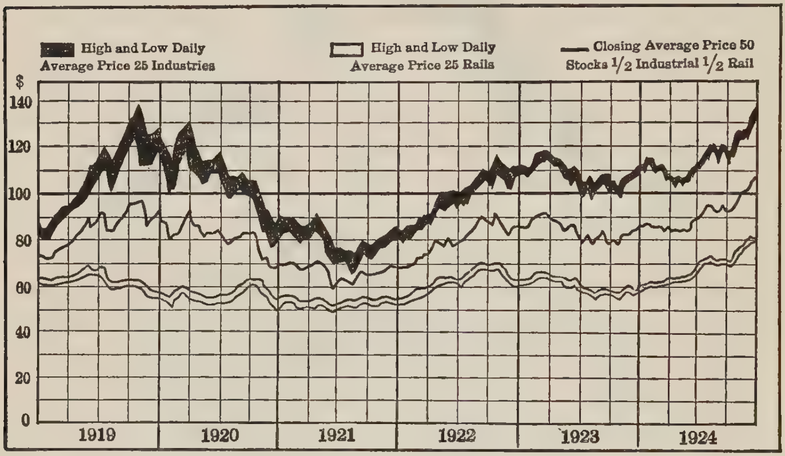


FIG. 47. — Average Prices of Stock on the New York Stock Exchange.[1]

The crisis or period of liquidation. The period of prosperity is one in which maladjustments in industry are prone to appear. By this we mean that certain industries become developed beyond the existing needs of the market, which means in turn that, while the quantity of goods which they are equipped to produce could probably be sold at some price, that quantity cannot be sold at a profitable price. Bearing in mind the fact that production is for a future market, whose absorbing ability can only be estimated, and the further fact that the period of prosperity is characterized by a magnificent optimism and forgetfulness of the chastisements of the past, one can easily understand the difficulty of securing an accurate adjustment of productive ability to consuming ability. Some of the major lines of industry are almost inevitably characterized by an increase in the output at a much more rapid rate than the increase in the demand for the goods.

---

[1] This and the succeeding figures are, unless otherwise indicated, taken from *Commerce Yearbook, 1924.*

As the era of prosperity draws near its close, uneasiness as to the future begins to manifest itself — the fear lest prices may not continue to rise. This leads some manufacturers to curtail or to postpone their projects of expansion and to decrease somewhat the volume of production. This in turn reacts upon the workers' control over goods and thus adds to the growing difficulty of disposing of goods at a profit. Sooner or later some manufacturers who have commitments to meet and who no longer can rely upon

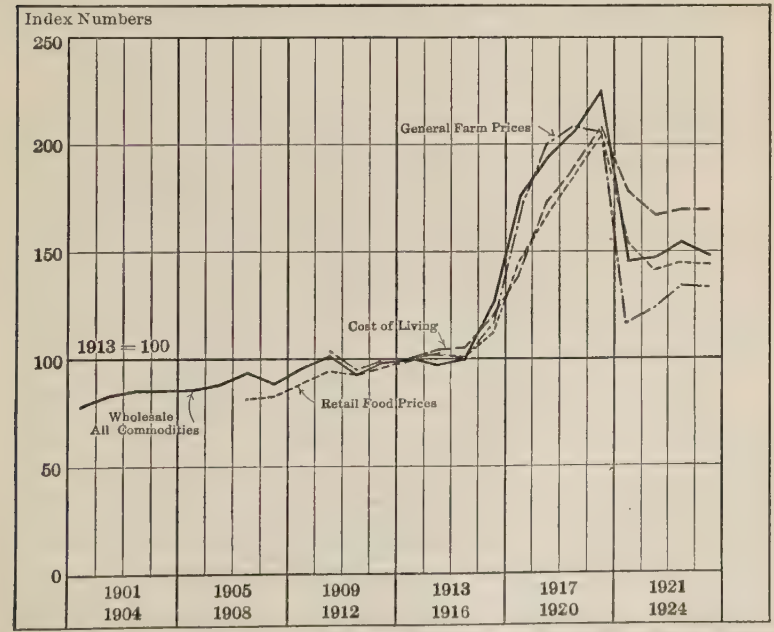


FIG. 48. — Comparative Annual Indexes of Wholesale, Retail, and Farm Prices.

the banks for financial assistance will be forced to sell goods for what they can get for them, and when once the drop in prices starts it becomes an avalanche. In place of the optimism of the previous era deepest pessimism reigns. No one knows how far prices will fall, and the one thought of every business man is to liquidate his affairs in the shortest time possible. The result is that each attempts to secure the settlement of claims which he has against others, and in this enforced liquidation prices are slashed. The behavior of wholesale prices in the period of crisis is graphically portrayed in Figure 48.

When prices break production falls off abruptly. This is not true of agricultural production, since farmers necessarily make commitments for several months in advance, but in manufacturing it is marked. In the chart below the annual production of pig iron for a number of years is given. Iron is directly or indirectly a raw material for every manufacturing industry, and the almost

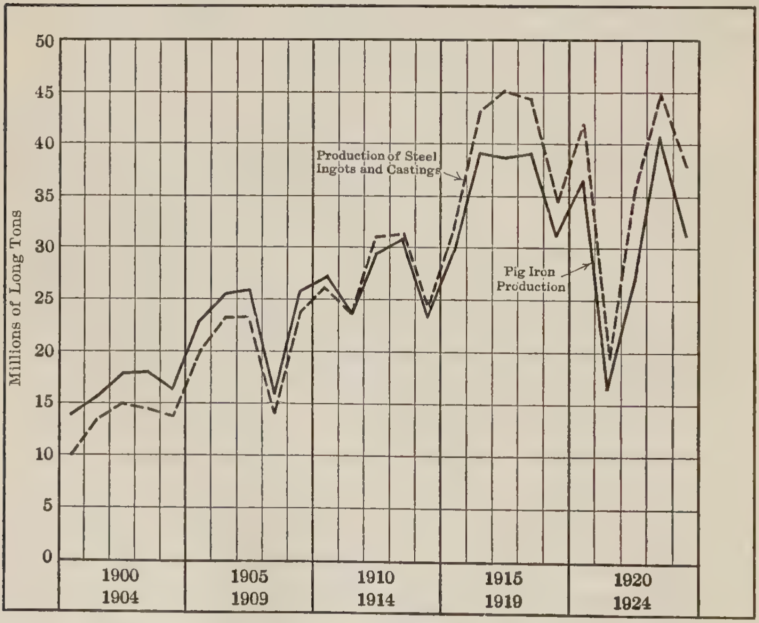


FIG. 49. — Production of Pig Iron and Steel Ingots and Castings in the United States.

universal curtailment of projects of extension is very sensitively recorded in the physical volume of production of pig iron.

The reaction of manufacturing in general to the period of crisis is shown in the following chart of manufacturing production and unfilled orders. There is a remarkable similarity in the trends of the three curves during the months of crisis. This very marked decrease in the physical volume of production is in large measure due to the change in the relation of the costs of production to the price of the product of the manufacturer. In this period the price of his product falls more rapidly than many of his costs of production — especially wages, interest, etc. — so that he may have to curtail materially his laboring force or even be compelled to close down his plant entirely. Unemployment therefore limits the consuming power of laborers, which in turn tends to force prices to a still lower level.

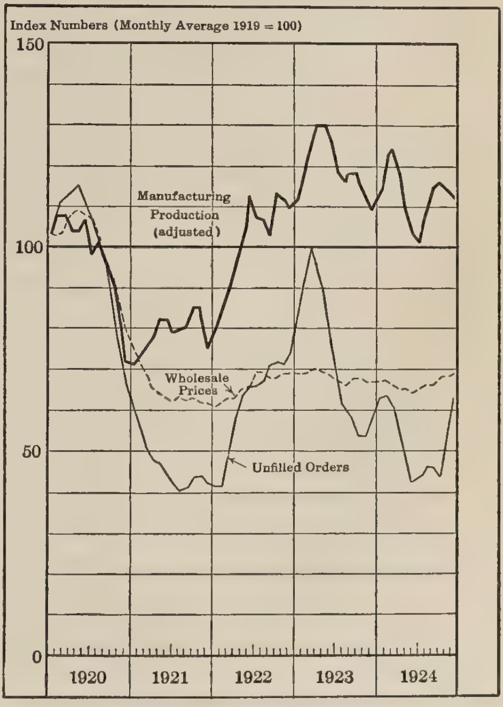


FIG. 50. — Indexes of Manufacturing Production, Unfilled Orders, and Prices.

In this period many business firms fail, and many others become so weakened in the effort to survive that they are unable to live through the ensuing period of depression. The chart on page 512, showing the aggregate liabilities of commercial firms which failed in the years 1917–1923, gives an illuminating picture of the connection between prices and profits and also of the cumulative effect of business failures. Credit relationships are nowadays so interlaced that the failure of one firm may bring about the failure of other firms in a continually widening circle.

During the period of crisis interest rates remain high, as will be seen by a glance at Figure 46, page 507, and the prices of stocks which have started to decline in the era of good times now come tumbling down.

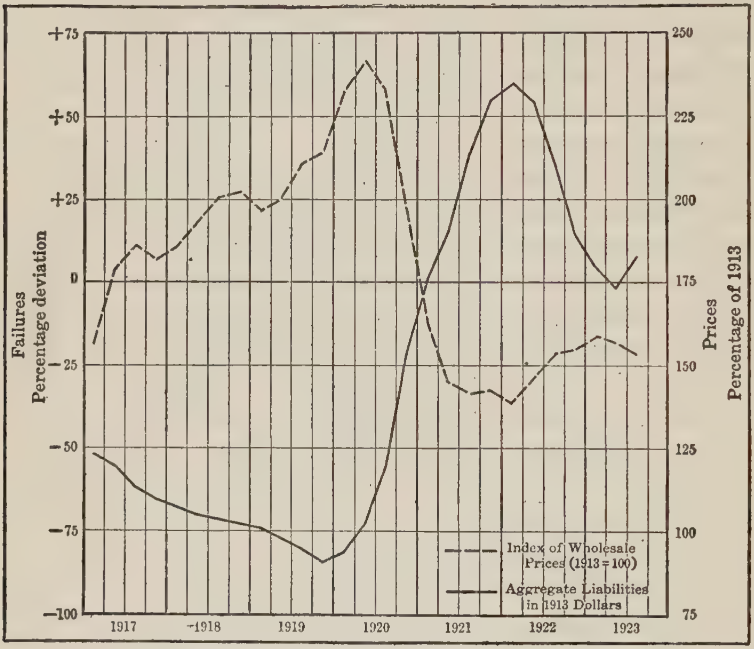


FIG. 51. — Index of Wholesale Prices and Cyclical Deviations from Normal of Liabilities of Commercial Failures in the United States.[1]

Panics. Not infrequently in the past a crisis has degenerated into a *panic*. Essentially a panic is the result of general and profound distrust of the banking and credit system. The panic of 1907, to illustrate, was precipitated by a run on a New York bank suspected of being involved in the failure of a prominent stock exchange firm. This bank was forced to close its doors. General loss of confidence in the banks caused runs on other banks, and some were forced to suspend payments while others adopted this measure for safety's sake. The wild desire to convert deposits into money compelled most banks to call in their loans to protect themselves against a possible run and soon forced them to restrict cash payments and to resort to other emergency expedients to conserve their cash resources. The money shortage became so acute that for several weeks money commanded a premium over checks on perfectly solvent banks. Interest rates, particularly on call loans, rose to exorbitant figures, and many speculators and others carrying securities by means of call loans were forced to sell out because of the difficulty and expense of renewing the loans. Prices on the Stock Exchange therefore collapsed. Failures multiplied, and in the general crash were involved firms which might have withstood the shock if they could have secured the credit accommodations which they merited. It was a case of absolute and unreasoning lack of confidence. Each suspected the solvency of his neighbor, hoarded his own cash, and refused assistance lest his own position be imperiled. But business relations are too closely intertwined for individual measures of self-protection to be of much avail in such times; in fact they intensify the danger. In the absence of any effective mechanism for coöperative action, the situation got out of hand and the panic swept the entire country.

---

[1] Derived from tables given in *Statistical Abstract of the United States, 1923*, p. 680.

The fact that the crisis of 1920 did not develop into one of the worst panics which the country has ever known can be ascribed to the improvements in our banking system which the establishment of the Federal Reserve System has brought about. As the reader is aware,[1] under the Federal Reserve System reserves are made mobile and emergency measures may be taken to prevent the loss of confidence which underlies panics, whereas under the banking system as it existed in 1907 it was a case of *sauve qui peut.*

The period of depression. The period of crisis — of rapid liquidation of credit obligations — is relatively short, usually lasting not more than a few months at the most. It is not impossible that it may be followed by an immediate recovery in business activity; a change in world conditions of trade, an unusually profitable harvest, or some like event may provide the tonic necessary to infuse life and vigor into business; but in the majority of cases depression seems to be the aftermath of crises.

---

[1] See Chapter XXIII.

The decline of prices continues during the depression but is more gradual than during the crisis. Wholesale prices and the prices of producers' equipment reach their low limits sooner than retail prices, which continue to fall until they are again in adjustment with the other prices. Labor is only partially employed, because producers are more interested in working off their accumulated stocks than in manufacturing new goods. On the other hand the process of liquidation has brought about a decline in the costs of production. Wages have declined, and labor has become more efficient through fear of unemployment. As bank loans are paid off and new loans avoided, bank reserves increase and the interest rate declines. In some cases the reorganization of a concern may wipe out a part of its fixed charges, or refunding operations may result in a reduction in the interest rate on long term obligations.

The period of depression then is one in which there is a complete rearrangement of internal and external relationships. Production necessarily becomes more efficient. The volume of production is kept at a low level until surplus stocks are disposed of, and new commitments are not made until there is a reasonable assurance of profits. In other words the period of depression, gloomy and unpleasing as it is, serves as a breathing spell for business; it is a period when readjustments in the whole complex structure can be made.

The recovery. In the course of time the stocks of goods available for immediate consumption approach exhaustion and prices rise in response to the renewed demand. Rising prices at the outset are not general; they occur first in one line and then in another, but gradually the movement extends. The situation is favorable to a renewal of activity — low wages and efficient labor, low rates of interest, relatively small stocks of goods, etc. — and the period of recovery may give way to the period of prosperity, when rising prices and increasing profits in all lines become manifest.

But the period of recovery does not always usher in a renewal of prosperity; if the conditions are not favorable to prosperity, recovery may be followed by a period, lasting possibly for several years, of comparative equilibrium. Such a period followed the depression and recovery of 1908. It is for this reason that we consider the business cycle to begin with the period of prosperity and to end with the period of recovery. While one cycle often follows directly on the heels of the other, business is not always undergoing cyclical fluctuations.

Motivating forces in the business cycle. Our examination of the sequence of events in the typical business cycle has made it clear that each phase normally has within itself the seeds of the next. The crisis develops logically from the conditions generated during the period of prosperity and is in turn succeeded logically by the periods of depression and recovery. But it does not seem to be true that the period of recovery brings about conditions which necessarily start a new cycle with a renewal of intense activity. Once started, the business cycle seems to be self-generating, but for the initial impetus which starts it going we must look to outside causes.

We are here face to face with one of the unsolved problems of economic science, a problem which is at present engaging the attention of many able economists. While a mass of interesting facts and relations has already resulted from this research, some of which have been drawn upon to illustrate the earlier part of this chapter, and while some very interesting hypotheses have been announced, it is safe to say that the complete solution of the problem is yet to be found. The reader will find other books upon the subject which will repay his attention. Its further investigation however lies outside the scope of this elementary text.

Effects of fluctuation. The fluctuations which attend the business cycle are generally injurious to all concerned. During the period of prosperity and rising prices, as we have seen, the rate of wages of the laborer rises less rapidly than the prices of the commodities which he consumes. This adverse influence is partly counteracted by the greater regularity of employment and the opportunity to work overtime. While for any given laborer the net result may or may not be advantageous, labor as a whole is apt to enjoy at this period a greater command over economic goods than at any other phase of the cycle.

In the period of crisis wages do not fall as rapidly as the prices of goods, and therefore the worker who can hold his job gains in real wages. However the periods of crisis and depression are marked by widespread unemployment. It has for example been estimated that the number of employed wage earners in the United States declined by more than five millions between the third quarter of 1920 and the first quarter of 1922. According to the figures of the United States Bureau of Labor Statistics, employment in manufacturing industries declined by about thirty-five per cent from June of 1920 to January, 1921. Crisis and depression bring a serious curtailment in the scale of living of many and to an appreciable number actual privation. The laboring class as a whole is a heavy loser during this period. Not only is the laborer the chief victim of the fluctuations in business activity; he is utterly helpless to remedy conditions.

It is particularly difficult to estimate the effects of the business cycle on the entrepreneurs. It may be argued with a certain degree of plausibility that the intoxication of the period of prosperity is worth the after effects of crisis and depression, but from what the reader already knows of the way in which individuals evaluate future income he will be able to appreciate the fact that an income spread evenly over, let us say, a period of ten years has greater utility than an income which is very large the first three years, negative in quantity the fourth year, very small during the fifth and sixth years, and moderate during the remaining years of the decade, even if in the second case the total income of the decade be somewhat greater than the total income in the first case. In general the entrepreneur stands to gain by a reduction in the fluctuations of business activity.

Debtors and creditors, as well as those with fixed incomes, are affected by cyclical fluctuations in prices as they are by price movements in general.[1] A man who contracted a loan in 1913 to be repaid early in 1920 gave back at the time of payment much less in actual purchasing power than he had received, and because of the increased money income which he probably enjoyed in 1920 the repayment of the loan occasioned less of a sacrifice than the same payment would have caused in 1913. On the other hand the creditor received a much smaller amount measured in terms of commodities and services than he parted with in 1913.

Similarly one with a fixed income suffers a reduction in purchasing power during the period of rising prices and gains correspondingly when prices fall. These fluctuations introduce an element of uncertainty which has the effect of reducing the total utility of his income.

We should not neglect to emphasize the effect of fluctuations upon society as a whole. Periods of rising prices foster the expansion of the physical equipment of some lines of industry beyond the point of profitable return. Some excess capacity in plants is necessary to meet the " peak " load, but to provide capacity now for a demand which will not materialize for ten or fifteen years is, in the majority of cases, an uneconomical expenditure. Few business men would do this intentionally, but in the speculative frenzy of extreme prosperity it seems impossible for all correctly to evaluate the future conditions, and the result is over-expansion.

On the other hand the spectacle of millions of men out of employment for months or years, not because there is no work in the world which can be done, but because the mechanism which controls their activity has broken down presents, from any point of view, an unnecessary waste of man power in a world which is not over-abundantly supplied even with the necessities of life.

From the standpoint of the individual members of society and from that of society as a whole, there is much to gain and little to lose by a stabilization of business or a reduction in the range of business fluctuations. Perhaps that may never be perfectly accomplished, but if business fluctuations can in some measure be controlled society will have achieved the partial solution of one of the most vital and baffling problems of our modern life.

---

[1] See Chapter XXIV.

Remedies for business fluctuations: control of credit. Among the many proposals advanced as remedies for the cyclical fluctuations of business three are of especial interest. The first concerns the control of credit. We have seen how the extension of credit provides the necessary purchasing power for industrial expansion of all sorts. In the ordinary course of events credit continues to be extended until prudence or legal requirements compel the bankers to adopt a sterner policy. This curtailment of credit may in itself be sufficient to force liquidation and to precipitate the crisis. But before this point has been reached there has been time for business activity to get out of balance. The question immediately presents itself as to whether it would not be feasible and desirable for credit extensions to be curtailed at some earlier point. Business activity might not be so intense, but on the other hand the reaction would not be so disastrous.

To make any such policy effective the control of credit would have to be lodged in the hands of an organization able to take a broad view of the situation and animated with a spirit somewhat above that of immediate profits. In Great Britain the Bank of England is in a strategic position, by virtue of its divorce from politics and its leadership in banking, to assert this control. In the United States the Federal Reserve Board and the several federal reserve banks should be able to exercise this function, but they were too tied up with the financing of the government war loans to act energetically in preventing over-extension of credit during the period of prosperity preceding the crisis of 1920, and it remains to be seen whether they can exercise in the future a sufficiently effective control over the extension of credit by member banks to slow down the speculative activity which is the breeder of crises.

2 **Forecasting.** The second line of attack is in the direction of determining the trend of events and then taking such precautionary measures that the predicted event (if it be unfavorable) shall not

come to pass. Forecasting involves the prediction of the course of future events on the basis of the experience of the past. It is predicated on an accurate knowledge and just evaluation of the relative importance of each event in the business cycle. In the light of this knowledge of the past, present events and trends can be evaluated, and by taking proper precautions over-expansion of business can be avoided.

As yet forecasting, despite its progress in recent years, is in its infancy, and it is little used by business men in the conduct of their practical affairs. But it is not too much to anticipate that, with the improvements of economic science and the proof of its value in the conduct of business affairs, entrepreneurs will come more and more to rely upon forecasting services and the extreme cyclical fluctuations will in some degree be thus eliminated from business.

3 **Stabilizing the monetary standard.** Since fluctuations in the quantity of gold are among the most fundamental of the causes of price changes, it follows that any plan which would provide a more stable monetary system than the present one based on the gold standard would tend to reduce the fluctuations of the business cycle. It is this that gives special pertinence in this connection to the movement for a stabilized dollar to which our attention has already been directed in the preceding chapter.

CHAPTER XXVI

INTERREGIONAL TRADE

Division of labor, trade, and transportation. Of the predominant place of division of labor in the modern economic organization the reader is now fully aware. Neither the individual nor the group is self-sufficient; each individual expects to purchase from others the major portion of the goods he consumes, and each community relies in part upon the productivity of other communities for the satisfaction of its wants. Indeed it appears that the most prosperous communities are the very ones which are least self-sufficient and that the economic progress of mankind has generally involved an increasing loss of local as well as personal self-sufficiency. In a society which has embraced the institutions of private property and personal freedom all division of labor is conditioned upon trade, while territorial division of labor is conditioned also upon transportation. These truths, which may have appeared almost too self-evident to require statement, will serve to introduce us to certain subjects of great moment, embracing economic principles by no means at once self-evident and in truth cumbered with certain popular fallacies which are not to be lightly brushed aside.

Trade, in its simplest form, consists merely in the exchange of goods or services in the same place. Most trade however involves more; namely, the exchange of goods between different places. It is this sort of trade which presents the most interesting and important economic problems and, since the principles of the simpler form of trade are included in those of the more complex, we are justified in confining our further study of trade to the subject of interregional and international trade.

The bases of interregional trade. So many popular fallacies have clustered about the problem of interregional trade that it is essential to inquire somewhat minutely into the nature of the gain that comes from regional division of labor and interregional trade. The trade in some articles is readily enough explained by the fact that in no other way could the community possibly obtain these particular goods. Diamonds, emeralds, and certain other precious stones could not be enjoyed by Americans if they were not brought in through trade. The residents of a great industrial city could not have fresh fruits and vegetables save for the trade of the agricultural hinterland.

The real problem of trade however has to do with the many commodities which apparently might be produced at home and which are nevertheless regularly brought in from other regions. The people of New York State could make their own automobiles instead of leaning so heavily upon Michigan; the United States could produce all the sugar it needs instead of only a part. In fact the literal impossibility of producing an article at home can seldom be cited as the reason for its import. The real reason is that by calling upon the resources and industry of other communities goods may be obtained at lower cost than would be incurred in their production at home. Each region thus looks for the things it needs to those parts of the country or of the world where each commodity is produced most cheaply and conversely finds the rest of the country and of the world calling upon it for those particular commodities which it produces at lowest cost. So each region, because of trade, is enabled to devote its resources and energies to the lines of production for which it is best adapted as evidenced by its ability to produce at low cost.

Local advantages and their causes. Many factors combine to determine whether one region can produce a good more cheaply than another. The rich soil and favorable climate of Cuba make it possible to grow sugar at a lower cost than in the United States; the exceptional quality of the manganese ore of the Caucasus Mountains, India, and Brazil gives those regions such an advantage in the extraction of that product that virtually all the manganese used in the steel industry of the United States is imported, in spite of the fact that we ourselves have large reserves of low-grade manganiferous iron ore. In other cases the low cost of production is due to the economies which arise from large scale production, the extensive use of machinery, and the minute subdivision of operations. The great reserves of iron and coal of England, together with an intelligent and industrious population, large accumulations of capital for investment, and the application of large scale methods of production, have given the British advantages in the manufacture of many articles. Numerous other illustrations of the way in which the products of a region have been determined by superiority in natural resources, by peculiar skill on the part of its laborers, by relatively great accumulations of capital, or by economies in the organization of industry will at once occur to the reader. We need only point to the fact that differences in productive advantages are enjoyed by different regions and that these lead to regional specialization in production and the exchange of the surplus products through trade.

The gain from specialization and trade. It is doubtless evident from what has been established to this point that specialization by each region in the production of the goods for which it is best fitted must increase enormously the total of the world's production and correspondingly add to the goods available for the satisfaction of human wants, but the situation will be clearer and the causal relation brought more out into the open if we inquire somewhat minutely into the economic effects of interregional trade. For this purpose we may set up a hypothetical case which, in spite of its obvious artificiality, will serve to expose the principle in which we are interested.

Let us suppose that country A and country B can produce either wheat or corn, but that A has an advantage in the production of wheat, while B has an advantage in the production of corn. To make the example concrete, let us say that with the expenditure of a unit of labor and capital (composed of so many hours of labor, so much machinery, fertilizer, etc.) A can produce either two bushels of wheat or one bushel of corn, while B with the same expenditure of effort (one unit of labor and capital) can produce either one bushel of wheat or two bushels of corn. Then if A and

B each spend two units of labor and capital and each produces both corn and wheat the result will be:

A	2 bu. of wheat + 1 bu. of corn
B	1 bu. of wheat + 2 bu. of corn
Total	3 bu. of wheat + 3 bu. of corn

If each country spends its two units in the production of the good in which it has a real advantage the result will be as follows:

A	4 bu. of wheat
B	4 bu. of corn
Total	4 bu. of wheat + 4 bu. of corn

In this case there is a gain in total product of one bushel of wheat and one bushel of corn, but neither country wants to be without the other grain. Rather than do without corn entirely, A would be willing to spend one unit in producing a bushel. Let us then further assume that A exchanges wheat for corn with B at the rate of one bushel of wheat for one bushel of corn.[1] If two bushels are sold, each country would have wheat and corn, as follows:

A	2 bushels of wheat + 2 bushels of corn
B	2 bushels of wheat + 2 bushels of corn

Comparing this table with the first one we see that A has gained a bushel of corn and B has gained a bushel of wheat, as the result of specializing in production and exchanging the surplus.

It may however happen that a country finds itself in the unhappy position of having no natural advantages. Its soil is sterile, nature has not endowed it with any mineral resources of prime importance, and its population is lethargic; yet we find the country has an export trade. On the other hand there are countries which are so favorably situated that they have advantages in many different lines of production; yet they regularly import certain goods from other countries which are actually inferior to them in those lines. This situation, as to the practical truth of which there can be no possible doubt, may seem anomalous if our previous example has led us to the erroneous conclusion that each country will necessarily produce for itself everything in the pro-

[1] In these illustrations we ignore the transportation charges.

duction of which it has an advantage, with the corollary that the country which has no superiority in any line can have no production at all.

To explain the situation let us assume two other countries, C and D. Country C with one unit of labor and capital can produce either fifteen yards of cloth or one ton of steel. Country D with one unit of labor and capital can produce ten yards of cloth or one half a ton of steel. C has an advantage in the production of both cloth and steel, but its advantage in the production of steel is greater. Compared with C, D has no advantage in the production of either commodity, but its "least comparative disadvantage" lies in the production of cloth. If each country devotes one unit of labor and capital to the production of cloth and two units of labor and capital to the production of steel, the amounts produced by each will be as follows:

C with 3 units	15 yards of cloth + 2 tons of steel
D with 3 units	10 yards of cloth + 1 ton of steel
Total	25 yards of cloth + 3 tons of steel

If however with the same expenditure of effort as before, C specializes on the production of steel, and D on the production of cloth, the amounts will be as follows:

C with 3 units	3 tons of steel
D with 3 units	30 yards of cloth
Total	30 yards of cloth + 3 tons of steel

From the standpoint of total production specialization has brought about a gain of five yards of cloth, but this is of no particular interest to either country unless the terms of exchange are such that each country will have a total greater than in the former case. C would be unwilling to sell steel for less than fifteen yards of cloth a ton, since it would be more profitable to manufacture the cloth at home; D, for the same reason, would be unwilling to pay more than twenty yards of cloth for a ton of steel. A rate of eighteen yards of cloth for one ton of steel falls within these limits, and we will assume that this is fixed and that C sells D a ton of steel for eighteen yards of cloth. Then we find, after the exchange has been effected, the following situation:

C	18 yards of cloth + 2 tons of steel
D	12 yards of cloth + 1 ton of steel

A comparison of this with the first table will show that C has gained three yards of cloth and D has gained two yards of cloth. In spite of C's greater productive advantage in both industries, it is worth its while to specialize and exchange products with D. Of course the people of D still have less to show for their efforts than the people of C, but they would have still less if they were unable to specialize on cloth and get their steel by exchange.

Conclusions. From this example we may draw an important generalization. It is to the advantage of a country or region to devote its energies, not to all the lines of production in which it may have superiority, but to those in which its superiority is greatest, provided trade gives it the opportunity to secure other goods from abroad. Likewise the country which has no line of superiority will find it advantageous to devote its energies to those lines in which its inferiority is least marked, provided the opportunity of trade with other regions is open to it. This generalization is called the *law of comparative advantages* and is stated in formal language as follows: *A country gains by concentrating its energies in the production of those goods in which it has the greatest comparative advantages or in which it has the least comparative disadvantages.*

Our conclusion is that regional specialization and interregional trade promote economy in the use of the productive forces of society. By utilizing the most fertile soils and the richest mines, by turning labor and capital into the industries for which each region is best suited, society as a whole has a greater quantity of consumable goods to share among its members, and each region is enabled to secure a greater quantity of consumable goods than otherwise would be possible.

Since the gain from trade consists in the increased command over consumable goods it follows that imports are the aim and end of commerce. We export in order that we may import. This conclusion is somewhat at variance with popular belief, but if it be true that commerce exists to serve men's wants, its benefits are to be measured by the command it gives over consumable goods. Our exports are of importance only in so far as they provide an easier means of securing the goods imported than by making them at home, and the greater the quantity of imports which we can obtain with a given quantity of exports, the more fortunate our position and the greater the national income. A falling off in the export trade is to be considered just cause for public alarm only because it may indicate that our buying power in foreign markets has decreased and therefore our command over foreign goods is less than formerly.

These conclusions were derived without reference to political boundaries, and they apply as well to regional specialization within the confines of a single country as to specialization between countries which are politically distinct. The gain which comes to the United States from purchasing the fine woollens of England with wheat is of exactly the same character as that which the people of New England experience when they buy western foodstuffs with their manufactured goods.

Diminishing returns and the law of comparative advantage. In every country hundreds and thousands of different articles are produced. Why, if the law of comparative advantages really holds good, should not each country, or each section of each country, concentrate on the production of one article only, and secure all the other goods it needs by exchange?

In partial answer to this question it may be said that the existence of a comparative advantage in the production of any commodity depends to some extent upon the quantity produced. There is a limit to the economies of large scale production; sooner or later all industry operates under conditions of increasing cost. A rising marginal cost of production is an indication that the comparative advantage in the production of additional units of the good is becoming less. The United States for example has such decided advantages in the production of wheat that it regularly has a surplus of several hundred millions of bushels for export, but if it were to attempt to supply the entire world with wheat the cost per bushel would rise far above that which would obtain in at least some of the countries which now import wheat. Rather than pay the price which in such circumstances would be necessary to remunerate the American farmers for their efforts, foreigners would grow wheat at home. Even now England, the great manufacturing nation, finds it profitable to produce a portion of the wheat she consumes rather than to buy it all from America or other nations. The law of diminishing returns therefore sets limits to specialization, and we find few products of any consequence in the world's commerce which are produced in one region exclusively, just as we find few important regions with only one industry.

Transportation and the law of comparative advantage. To complete the answer to the question raised above we must also take into account the factor of transportation. In our mathematical examples we ignored the costs of transportation, but in actual life they offer one of the greatest obstacles to the development of regional specialization and trade. The reason that the merchants of any region import a commodity from outside is because it can be sold in the domestic market at a price which is lower than the price at which a similar article of domestic manufacture can be sold, and in determining the sale price the cost of transportation must be included. English manufacturers may be able to produce certain grades of bricks at a lower cost than can American manufacturers, but if the cost of shipping them to New York is so great as to raise their price above that of similar domestic bricks, the advantage of the English manufacturers so far as the New York market is concerned is without significance. There are many goods of which this is true, goods which are commonly produced in almost every locality, because the cost of transporting them from one place to another is so great as to wipe out entirely the advantage which any one locality may have. Nearness to market therefore gives an advantage in production which may outweigh positive disadvantages, as is shown for example by the fact that common bricks are almost everywhere a local product, seldom brought in from any great distance even though other places may enjoy advantages in their manufacture.

The control of trade. Every producer of goods naturally seeks to sell his product in the market where he obtains the highest price, be it near at hand or far away. The consumers just as naturally endeavor to make their purchases where goods may be obtained most cheaply. All the agencies of trade and transportation are enlisted in the effort to aid producers and consumers in the accomplishment of their respective purposes; every class of middlemen is actuated by the same motives as the other producers and the consumers to buy in the cheap market and sell in the dear market. There is a tendency therefore for goods to move from where their prices are low to where their prices are high. By thus reducing the supply in one market and increasing it in another, the price in the one is raised and in the other is lowered, till the prices in different markets are brought into equilibrium.

Equilibrium of prices does not mean equality of prices. The movement of any commodity from the low priced to the high priced market reaches a limit when, as the result of such movement, the difference between the two prices is brought down to the cost of shipping the commodity from the first market to the second. Thereupon the movement ceases, price equilibrium having been reached. If transportation could take place instantaneously and without cost we would find the price of each commodity the same in all markets. As it is equilibrium is attained when the prices of a commodity do not differ by more than the cost of transportation. This principle may be stated in the following law: *In the absence of restraints on trade, the prices of any commodity in two markets tend to differ by an amount not greater than the cost of transporting the commodity from the market of lower price to that of higher price.*

Thus price and the cost of transportation control the flow of goods from one place to another. There will be trade so long as there would otherwise be a difference in the prices of the same commodity in two markets greater than the cost of transportation between them. Where the difference is less than this there is no interregional trade, as is the case of many commodities of which the supplies are relatively local; bricks, for example. The ability to specialize to advantage is evidenced by a relatively low cost of production, which enables the producers of one region to undersell those of other regions, and the opportunities for the extension of regional specialization and trade are brought to light by differences in the prices of the commodity in the various markets. Under competitive conditions price serves as an automatic control over regional division of labor and interregional trade.

Obstacles to trade and price equilibrium. The restraints which may prevent trade and so prevent the attainment of price equilibrium are numerous. Frequent examples have already been encountered in our study, and we shall later give attention to one of the most important, the customs tariff. In the present connection however we must take special note of one commodity (if this term may be so used) which is peculiarly susceptible to obstacles restraining movement in response to inequalities of price. This is labor, whose price, wages, may vary widely and permanently between different countries and different parts of the same country without being corrected by the migration of the laboring populations. Domestic servants in one part of the world may be obtained at an extremely low figure, while their services in another country command a high price. Labor of a particular type is paid different rates in different sections of the country, and there is often material difference between the wage scales of the city and the immediately surrounding countryside. In some cases the difference may be only nominal, for though the wages of the carpenter in the village may be lower than those obtainable for the same grade of skill in the city, his living expenses may be lower also, so that his real wages — the command over commodities which his money wages give him — are as high as those of the city laborer. But always there remain real differences which are not thus to be explained away. Labor as a matter of fact is a relatively immobile "commodity." The personal element is always present. A laborer who has established himself and his family in a particular environment dislikes intensely to break the old ties and form entirely new associations elsewhere, and he will do so only when he becomes convinced that the benefits derived from the new location are of a very substantial nature. Wheat will be sent back and forth from one city or one country to another as the price fluctuates; labor moves much more slowly and reluctantly.

The domestic trade of the United States. From the foregoing study of general principles we may now profitably turn to a picture of the actual commerce of the United States, which falls naturally into two parts, domestic and foreign. It is noticeable that in most discussions of trade attention is apt to be directed predominantly to the country's foreign trade, so much so that many persons have an entirely distorted notion of the relative importance of the two branches of national trade. As a matter of fact the domestic trade of the United States is much larger in volume and hence presumably more significant to the welfare of the nation than the foreign trade. Unfortunately there are no statistics of domestic trade, as it is obviously impossible to require detailed records of transactions without seriously hampering trade. Exports and imports on the contrary are carefully registered at the custom-houses. Figures on the volume of foreign commerce are therefore available and are frequently cited in the newspapers and financial journals, the course of our foreign trade is eagerly followed by business men, and as a result popular emphasis is wrongly placed.

In the endeavor to make reasonable estimates of the volume of domestic trade many ingenious methods have been utilized, but it should be remembered that all figures of domestic trade are but estimates and at best subject to a wide margin of error. We do not need to discuss the methods used in determining the relative values of foreign and domestic trade, but merely to state the results. Depending on the method used, the foreign trade is variously estimated as amounting to from two to nineteen per cent of the domestic trade. In other words, taking the lowest estimate of our domestic trade, it is five times the foreign trade in volume; taking the highest figures, it amounts to fifty times the foreign trade. Whichever of these estimates is the more nearly correct, we conclude that domestic trade is unquestionably the more important. The popular exaggeration of the importance of foreign trade, particularly of the export trade, is based on misinformation as to the facts and misunderstanding of the principles of trade.

The foreign trade of the United States. This is not the place for an extensive survey of the foreign commerce of the United States, but the student of elementary economics should know something of its size and its general character. The table below gives the total value of exports and imports for a number of years and will serve to illustrate the phenomenal growth of our foreign commerce as a whole.

FOREIGN COMMERCE OF THE UNITED STATES
(in millions of dollars)

	Exports (domestic and foreign)	*Imports*
1800	70.9	91.0
1820	69.7	74.4
1840	123.7	98.2
1860	333.6	353.6
1880	835.6	667.9
1900	1,394.5	849.9
1910	1,744.9	1,556.9
1920	8,228.0	5,278.5
1923	4,167.5	3,792.0
1924	4,590.9	3,609.9

That the growth has been real and significant is self-evident, but the student should be cautioned against drawing hasty conclusions from the table. It is not true that between the years 1820 and 1920 the physical volume of exports increased by something over ten thousand per cent, as a comparison of these figures might seem to indicate, for during those hundred years there were various changes in the price levels, and the price level in 1920 was much higher than in 1820. The Bureau of Foreign and Domestic Trade of the Department of Commerce has made a comparison of the physical volume of trade by translating the values of the goods exported and imported in various years into 1923 prices. Thus the value of the average annual exports of domestic goods for the period 1910–1914 as reported was $2,130,000,000; in terms of 1923 prices it was $3,669,000,000, an increase due to the fact that 1923 prices were considerably higher than those of 1910–1914. Since the value of the domestic exports in

1923 was $4,091,000,000, it appears that the increase in the volume of exports was not ninety-one per cent, as would appear from a comparison of the first figure for 1910–1914 with the exports of 1923, but really about ten per cent. It is quite possible for the physical volume of trade to decline from one year to another, while the value of the trade increases. Professor Fisher gives the following interesting example of this:[1] "The money value of car wheels exported from the United States in one month was $12,000 and in a later month $15,000, from which fact we might infer that the quantity of these exports had increased. But the *number* of car wheels exported in the first of these two months was 2,200, and in the second only 2,100, showing a decrease. The price had increased faster than the number had decreased."

The commodity character of our foreign trade. The following table gives for selected years the commodity character of our foreign trade.

COMMODITY CHARACTER OF EXPORTS AND IMPORTS IN PERCENTAGES OF TOTAL EXPORTS AND IMPORTS

Commodity	*Imports*		*Domestic Exports*	
	1860*	1924**	1860*	1924**
Crude materials for use in manufacturing	11%	34%	68%	29%
Foodstuffs in crude condition, and food animals	13%	12%	4%	9%
Foodstuffs wholly or partly manufactured	17%	15%	12%	13%
Manufactures for further use in manufacturing	10%	18%	4%	14%
Manufactures ready for consumption	49%	21%	11%	35%
Total (in millions)	$354	$3,113	$316	$3,765

* Year ending June 30.
** Calendar year.

These figures tell the story of a revolutionary change in the character of the foreign trade of the United States, indicating the growth in relative importance of the nation's manufacturing industries. In 1860 we exported seventy-two per cent of our goods in crude form. In 1924 all but thirty-eight per cent of our domestic exports were partly or wholly manufactured, and the percentage of consumable manufactures had increased from eleven to thirty-five. On the import side the converse is seen to hold true. The percentage of manufactures ready for consumption which we imported declined from forty-nine in 1860 to twenty-one in 1924, the percentage of goods partly fabricated but requiring further processes increased, and the percentage of raw materials imported increased greatly.

---

[1] *Elementary Principles of Economics*, page 19.

CHAPTER XXVII

FOREIGN EXCHANGE

Payments between persons in the same place are made by means of money or checks, and the process has been sufficiently investigated in the previous chapters. Payments between persons in different places on the other hand involve certain special problems, owing to the necessity of making currency paid in one place available in another place. If in addition the two persons are in different countries, a further complexity may arise owing to the difference in monetary units. Since all the economic problems involved in payments between places in the same country are included in the more complex subject of international payments, the whole matter may properly be discussed under the subject of foreign exchange. We shall find it convenient in this discussion to assume conditions as they were before the World War, in order that we may avoid the abnormal complications caused by the monetary disturbances of the war and the immediately succeeding years.

An export of cotton. The simplest approach to this subject is by way of the study of typical examples of export and import transactions. Let us therefore assume that Richards & Green of New York has sold Harris & Company, Ltd., of London a certain number of bales of cotton for £1,000. Now it is evident that the ordinary forms of currency used in domestic trade are not available for payment in this case. It is within the realm of possibility that Harris & Company, Ltd., might effect payment by shipping a quantity of gold to Richards & Green. Gold is a commodity which has a ready market, and were Richards & Green to receive one thousand gold sovereigns it could sell them to bullion brokers or take them to the mint and have them coined into American dollars. But the shipment of gold is a troublesome and expensive operation which is practically never used for commercial payments. The universal currency for international payments is the bill of exchange.

The bill of exchange. We may assume, what would almost always be true in such a case, that the terms of the sale have called for a bill of exchange to be drawn by Richards & Green on Harris & Company, Ltd. The bill will be in the following form:

No. 726 New York, December 29, 1924
Pay at sight to the order of the New York Bank
..........One Thousand Pounds Sterling..........
(signed) Richards & Green
To Harris & Company, Ltd.
London, Eng.

The bill of exchange is an instrument of such importance that we must pause here to formulate its precise technical definition, which is as follows: *A bill of exchange is an unconditional order in writing, signed by one person (called the drawer) directing a second person (called the drawee) to pay a third person (called the payee) or his order a certain sum of money at some definite time in the future.* It will be noted that the example given above conforms with this definition. The bill of exchange appears in various other forms, of which the simplest is the ordinary bank check, with which the reader has become acquainted through the medium of an earlier chapter (Chapter XXII) and probably also from practical experience.

Financing exports. When Richards & Green ships the cotton it secures from the steamship company, or possibly from a railroad company if the shipment is to be made from an inland warehouse, a bill of lading, which not only shows that the company has received a certain number of bales of cotton and has agreed to transport them to England, but also serves as the evidence of ownership in the cotton. Anyone in possession of a properly endorsed bill of lading can secure the cotton from the steamship company, and it can be secured in no other way. Marine insurance must be taken out on the shipment in order that the owner may be protected if the cotton suffers loss or damage during the sea voyage, for ocean carriers are by law exempt from practically all responsibility with regard to the safety of shipments entrusted to them.

Richards & Green next takes the bill of exchange, with the two documents (bill of lading and insurance policy) and possibly others attached, to its bank, the New York Bank, for sale. The bank knows nothing about the reliability of Harris & Company, Ltd., but it knows that Richards & Green is an old, well-established house with a reputation for financial straightforwardness. If it buys the bill of exchange it will acquire a credit instrument bearing the signature of the responsible officer in the concern and having the same standing as a promissory note signed by the same officer. If Harris & Company, Ltd., fails to meet the bill when presented for payment the bank can enforce payment from Richards & Green. Furthermore the bank acquires possession of the bill of lading, which will give it the title to the cotton. Now cotton is a staple article, not perishable, and commanding at once a definite price in all the important markets of the world. If the worst should happen and the bank could not collect from Richards & Green, it could sell the cotton and still not lose much. Relying therefore on the responsibility of Richards & Green and on the cotton as collateral security, it purchases the bill of exchange at the rate prevailing in the market at the moment for bills of that character. The rate, let us say, happens to be $4.84 per pound sterling, and Richards & Green therefore receives $4,840 in cash. The exporting firm has thus received payment for the cotton in American dollars immediately after shipment and has no further interest in the transaction, save for its contingent liability.

The bank now sends the bill, with the documents still attached, to the London Bank, with which it has correspondent relations, with instructions that it be presented at once to Harris & Company, Ltd., for payment and that the proceeds be credited to the account of the New York Bank. Harris & Company, Ltd., pays the bill when presented and receives the bill of lading, which entitles it to claim the cotton from the agents of the steamship company.

We may summarize these stages in the financing of the export of cotton as follows: (1) Richards & Green draws a bill of exchange on Harris & Company, Ltd., for £1,000. (2) Richards & Green sells the bill of exchange (with documents attached) to the New York Bank for $4,840. (3) The New York Bank sends the bill and the documents to its correspondent in London, the London Bank. (4) The London Bank collects £1,000 from Harris & Company, Ltd., and surrenders the bill of lading. (5) The New York Bank's deposit account with the London Bank is credited by £1,000.

Financing imports. A similar example will aid us to visualize the several steps involved in the financing of an import. We assume that White & Company of New York has purchased £800 worth of cotton goods from R. J. Littlefield of London. It is again possible for White & Company to ship a sufficient quantity of gold to be worth £800 in London, or Littlefield might draw a bill of exchange on White & Company, thus reversing the process described in the previous section. But as a matter of fact trade practice before the World War had placed the responsibility of making payment, as well as of securing payment, on the American merchant. So in this instance White & Company is required to place Littlefield in possession of £800 in cash or its equivalent in London. This it can do by purchasing from its banker a bill of exchange such as the following, often called a demand or sight draft.

No. 918 New York, January 5, 1925.
At sight pay to the order of R. J. Littlefield
......Eight Hundred Pounds Sterling......
and charge the same to our account
(signed)
New York Bank

To
The London Bank

The New York Bank sells this bill of exchange to White & Company at the prevailing rate for bankers' demand drafts, which we will assume to be $4.87. The draft therefore costs White & Company $3,896. White & Company sends the draft by mail to Littlefield, who presents it at the London Bank and receives £800 in cash, which amount is debited to the account of the New York Bank.

The banks and the exchange market. The part which the bankers play is that of middlemen. It would indeed be possible for the importers to buy bills of exchange from exporters and to that extent eliminate the bankers, and in the past this has been done. We know that in the first decade of the nineteenth century American importers rarely paid by bankers' drafts; in almost all cases the bills remitted were those drawn on British purchasers by exporters of cotton, grain, or other products. Today virtually all bills are in the first instance either sold to or bought from the bankers. The banker has a wide circle of customers importing and exporting every known ware and dealing with all the markets of the world, and he is able to buy and sell drafts in all currencies and in all amounts at a moment's notice. It is patently much simpler for an importer who must remit a draft for £168 9s. 3d. to purchase it from a banker than to find an exporter who has a draft for this precise amount, for the banker can draw a draft in any denomination so long as he has a credit balance abroad sufficiently large to cover it, while the size of the exporter's draft is regulated by his export. An importer of goods from China can secure a draft payable in Shanghai or Hongkong with no difficulty, and furthermore he knows and the Chinese exporter knows that the risk of loss inherent in a banker's draft is much smaller than in a draft drawn by one commercial house on another. Because of these advantages it is not surprising to find the banker at present the focal point of the foreign exchange market. Only rarely would an importer who is required to remit a bill of exchange send anything but a banker's draft, or an importer who has a bill of sale look for a purchaser outside his bank.

Middlemen always expect to sell what they buy, and bankers do not differ from other middlemen in this respect. When they buy bills of exchange it is because they expect to be able to sell bills of exchange. We have seen that the result of the purchase by the New York Bank of the bill drawn by Richards & Green was to give this bank a credit with the London Bank of £1,000. It was the existence of a credit such as this which enabled the New York Bank to sell its own demand draft to White & Company. By the same token the sale of this draft to White & Company gave the bank cash wherewith to purchase other bills of other exporters. Every time a bank buys a bill payable in London, it reduces the supply of cash in its own vaults and acquires a deposit in England. It will not buy bills of exchange simply to pile up unused credits in the London Bank, for the low rate of interest which it might get on such a deposit would not compensate it for the loss of funds in New York, and furthermore so one-sided a business would soon drain all its cash resources and compel it to suspend operations. On the other hand every time a bank sells a bill payable in London, it acquires cash in its own vaults and reduces its deposits in England. The bank cannot long continue a one-sided business of this sort without entirely exhausting its London credit and so being compelled to suspend operations.

If however the bank regulates its purchases of bills and its sales of its own drafts in harmony with each other, it will create and utilize its foreign credits and expend and replenish its cash resources at home in approximately equal amounts from day to day. These balanced transactions will permit it to keep its balance with its foreign correspondent neither unduly large nor unduly small and likewise to avoid both a shortage of cash and a surplus of idle funds in its home office. The bank dealing in foreign exchange intends that its operations shall approximate this equilibrium, while making its profit in foreign exchange by selling at a slightly higher rate than the one at which it buys.

But it will be noted that the bankers must depend for equilibrium in these transactions on the demand and supply of bills arising from the currents of international commerce. These currents of commerce are variable. At certain seasons of the year the volume of exports from the United States is extraordinarily heavy; at other seasons American merchants are importing unusually large quantities of goods. It is quite possible that, on account of the transactions of a certain short period, business men in America may owe £50,000,000 on account of imports received and be owed only £25,000,000 on account of exports. By buying the bills of the exporters, American bankers can acquire credits in London of £25,000,000 and be in position to sell their own bills to this amount to the importers. But this will leave £25,000,000 of the debts of the American importers unpaid. What happens in the exchange market when so great a discrepancy arises between exports and imports? Are some of the importing merchants left to shift for themselves? On the other hand, if the trade balance happens to be the other way, are some of the exporters left with no means of obtaining payment through sale of their bills of exchange? Before answering these questions we must make some inquiry into the forces which determine the rates of exchange.

The mint par of exchange. Since nearly every nation has its own peculiar monetary unit, international payments generally involve translation of one currency into another. When two countries, like France and Italy, have the same unit (even though it goes by different names), the translation of the one into the other involves no problem, the ratio between them being unity. Even when the units are different, the problem is merely one of simple arithmetic; namely, the determination of the ratio between the pure gold contents of the respective units. This ratio is called the *mint par of exchange,* and its determination is a necessary preliminary step in the study of the rate of exchange between any two nations.

Between France and Italy, the mint par is 1; *i.e.*, one franc is equivalent to one lira. As an example of countries with unlike units, let us calculate the mint par between the United States and Great Britain. The sovereign of Great Britain contains $113.00116\frac{2}{3}$ grains of fine gold, and the gold dollar of the United States contains 23.22 grains of fine gold. The ratio is $\frac{113.00116\frac{2}{3},}{23.22}$ or 4.8665+. Hence the British gold sovereign contains 4.8665+ times as many grains of fine gold as the dollar, and the mint par of exchange as \$4.8665+ per pound sterling, meaning that if ten thousand gold sovereigns of full weight were taken to the United States mint they would contain sufficient fine gold to furnish the gold content of \$48,665. For our purposes we shall be sufficiently exact if we consider the mint par of exchange as \$4.86 per pound sterling. The franc contains 4.48035 grains of fine gold; hence par of exchange with France is 5.1826 francs per dollar (23.22 ÷ 4.48035). The lira of Italy, the peseta of Spain, and the francs of Switzerland and Belgium have the same fine gold content as the franc of France, and consequently the same mint par of exchange with the dollar. Mint pars between the dollar and the other standard gold coins of the world are determined by the same principles.[1]

The rate of exchange. As usually expressed, the *rate of exchange between two countries is the price paid in the home country for one unit of the money of the foreign country payable in that country.* In other words the rate of exchange is the price of foreign bills of exchange. For example the rate in New York on London is the amount of American money that must be paid in New York for a bill of exchange calling for one pound of English money payable in London.[2] The rate of exchange is not the same thing as the mint par of exchange. When the rate of exchange in the actual market stands at the mint par we say that exchange is "at par." But this rarely occurs, and when it does it is only by chance. The mint par does not determine the rate of exchange; its significance lies, as we shall see, in the fact that it sets limits to the fluctuations of the rates.

The rate of exchange is a price and, in common with competitive prices in general, is determined by demand and supply. The demand for bills of exchange comes primarily from the importers and others who are under the necessity of making payments in foreign countries. The supply comes primarily from the exporters and others who have money payable to them in foreign countries. The bankers are the middlemen who, buying and selling, bring together the forces of demand and supply. In accordance with the general laws of price, the greater the demand for bills, the higher the rate of exchange; the greater the supply, the lower the rate. In terms of international trade, when imports are large compared with exports, the rate of exchange tends to be high; when exports are relatively large, the rate tends to be low.

---

[1] The reader will recall that we are considering conditions as they were before the World War, when the currencies of these nations were still on the gold standard.

[2] Rates of exchange may equally well be expressed the other way round, as the amount of the foreign money payable in the foreign country which may be obtained in the home country for one unit of its money. Thus the rate of exchange on France is often stated as so many francs per dollar. This is not the usual method however.

The gold points. But the rate of exchange differs from most competitive prices in one important respect. It can fluctuate only within definite limits. Having risen to the upper limit, the rate will stand there, no matter how great the demand, till changed conditions send it down. Having fallen to the lower limit, it can go no lower and will eventually rise. To make this principle clear, let us revert to the situation where Americans owe £50,000,000 and are owed in return only £25,000,000. As the discrepancy between the volume of bills which importers and others want and the volume offered to the bankers for sale begins to be manifest, bankers will tend to charge a higher price for their own bills of exchange in order to protect their foreign balances; and individual bankers will bid against each other for the bills of exchange which are offered for sale by exporters, for each will want to get as large a share of the possible credits abroad as he can.

The rate of exchange will be forced up, and the higher rate will tend to discourage some who want to make payments, and perhaps encourage others to anticipate the drawing of bills of exchange. But if the forces at work are strong enough the bankers will soon find their foreign balances exhausted, with importers still trying to buy bills of exchange. In this situation the banker, if he is not to cease operations entirely, is forced to resort to other expedients to build up his foreign bank balance. He may borrow money from a London Bank and thus have a balance against which to draw, but in all probability he will export gold from the United States for this purpose.

Gold equivalent to one thousand pounds sterling can be purchased in New York for $4,860, and it will cost something to cover the shipment of the gold — to take care of crating, freight, insurance, and the like. Suppose that these expenses amount to two and a half cents per pound sterling; then the total cost of placing one thousand pounds in London will be $4,885. The banker can therefore offer to sell bills of exchange on London at $4.885 per pound, and the importers who are under obligations to make immediate payment must pay this rate. True they could ship the gold themselves, but the banks can do it much more cheaply, and competition among different banks forces them to do so. This rate we call the *gold exporting point,* determined by the cost in New York of transferring one gold sovereign to London. It is the highest point to which the rate of exchange in New York can rise, and at that rate the bankers will continue to furnish bills to all comers, no matter how great the demand.

In like manner we may determine the *gold importing point.* Let us suppose that the supply of sterling bills of exchange, that is, bills drawn on London in pounds sterling, increases enormously, without any corresponding change in the demand. As we have seen, the supply of bills of exchange arises when exporters and others having debts to collect draw bills of exchange against the foreign debtors and attempt to sell them to bankers, and a demand arises when importers and others having debts to pay to foreigners go to the bankers to buy bankers' drafts. In this case therefore bankers are likely to have large credit balances abroad for which there is no call, and if the situation continues they will refuse to buy bills except at rates which will permit them to transfer their foreign balances to replenish their dwindling domestic cash resources. The cost of shipping gold from England to America is slightly higher than in the other case, for in this instance the bankers lose interest during the time of shipment; whereas, when gold is shipped from New York, the bankers can sell drafts to go by the same steamer and so avoid tying up their funds. If we assume the cost of shipment to be three cents per pound sterling, bankers will be willing to buy bills of exchange at the rate of $4.83. With a draft of £1,000 a banker can obtain a thousand gold sovereigns, bring them to the United States, and sell them for $4,860. The cost of doing so being $30, he will recover $4,830, the amount he paid for the bill. This rate is known as the gold importing point, and at this rate bankers will be willing to purchase an unlimited quantity of bills of exchange, no matter how great the supply.

The gold points then are determined with respect to the mint par of exchange and fix the limits within which the rate of exchange may fluctuate. If the rate rises to the upper gold point, gold will be *exported;* if it falls to the lower gold point, gold will be *imported.* The rise or fall in the rate is determined by the relation of demand and supply to each other, which in turn rests on the relation between our obligations to foreigners and the obligations of foreigners to us. If these are just equal, the rate would presumably be at or very near the mint par of exchange, and there would be no tendency for gold to flow at all; if at any time one greatly outbalances the other, gold will flow out to create the necessary foreign bank balances, or gold will flow in as bankers find this the only way to utilize their foreign balances.

Indirect exchange. We have thus far examined the machinery by which are settled the obligations arising out of the trade between two countries, as though each country had thus to make a separate settlement with each other country, much as the banks made separate settlement with each other before the days of the clearing house. If this were the actual case, there would be almost continual shipping in and out of large quantities of gold, since it is scarcely conceivable that the import and export transactions between any two countries should exactly offset each other except as a rare coincidence. We should expect to see the United States for example receiving large shipments of gold from Great Britain while at the same time it was making heavy shipments to Brazil. Now as a matter of fact international trade is always conducted on the clearing house basis, in accordance with which all the obligations of the residents of any given country to residents of all other countries are offset against all claims against foreigners. There is no international clearing house, but the result is accomplished with remarkable certainty and smoothness through the medium of foreign exchange. For this purpose sterling exchange has come to be the international currency. For many years bills on London have been in all markets a most acceptable form of exchange. In the settlement of many of our obligations we are required to remit sterling bills of exchange in preference to dollar bills or even to bills drawn in the currency of the local government.

Let us assume three countries, A, B, and C, having the following trade relations with each other:

A		*C*	
exports to B	£250,000,000	exports to A	£100,000,000
imports from B	225,000,000	imports from A	75,000,000
	+ 25,000,000		+ 25,000,000
exports to C	75,000,000	exports to B	125,000,000
imports from C	100,000,000	imports from B	150,000,000
	− 25,000,000		− 25,000,000

B	
exports to A	£225,000,000
imports from A	250,000,000
	− 25,000,000
exports to C	150,000,000
imports from C	125,000,000
	+ 25,000,000

Were we to consider the relations of any two of these countries alone, it would appear that one was under obligation to ship gold to the other. But when we take the relations of the three together we see that no gold need be shipped at all.

Let us assume that bills drawn on B are acceptable in each of the other countries, as sterling exchange is today acceptable in the markets of the world.

A has credits in B of £25,000,000, which it does not need in settlement of its obligations to B, but it can use these credits to draw bills of exchange payable to creditors in C. The creditors in

C will sell these bills to local bankers in C who will forward them to B for collection. The bankers in C will then have balances in B against which bills can be drawn in settlement of the indebtedness of C to B.

B being the financial centre, its residents play a comparatively passive rôle in the settlement of international indebtedness. Its net debt of £25,000,000 to A will be settled by paying to banks at home (in B) the bills drawn by the exporters in A, and its credit balance against C will be obtained by collecting from its own banks the bills sent by the importers of C. The two items thus offset each other. Great Britain has long been in approximately this situation, her foreign business, both as regards receipts and payments, being settled at home through bills drawn on London by foreigners.

The resulting balances in the banks of B will be as follows:

Accounts of Bankers of A	
Credit balances	
from exports to B	£250,000,000
from exports to C	75,000,000
	£325,000,000
Drawn against	
for imports from B	£225,000,000
for imports from C	100,000,000
	£325,000,000

Accounts of Bankers of C	
Credit balances	
from exports to B	£125,000,000
from exports to A	100,000,000
	£225,000,000
Drawn against	
for imports from B	£150,000,000
for imports from A	75,000,000
	£225,000,000

The exporters of B have remitted to them a total of £375,000,000 in drafts, and the importers have drafts drawn against them of £375,000,000 which will exactly offset each other.

In this case the combined exports of each of the three countries exactly equals its combined imports, and by means of bills of exchange drawn on one of the countries the transfer of gold between countries is entirely avoided. Actually a country has relations with all the other countries of the world, and the problem of settling trade differences with any particular country may not be quite so simple as represented in this example. Furthermore the indebtedness of a country will never exactly equal the sum of the debts of other countries to it, and some gold will flow in or out to settle this difference. Nevertheless the principle still holds good that exports to any part of the world may pay for imports from any part of the world, and this reduces the flow of gold to the smallest possible amount, representing the net balance between all exports (or similar items) and all imports (or similar items).

A case of three cornered settlement, substantially like our illustration, is seen in the trade between Great Britain, the United States, and Brazil. We normally export more to Great Britain than we import from her; we import more from Brazil than we export to her; and Brazil imports more from Great Britain than she exports to her. Our exports to Great Britain serve to pay for our imports from Brazil, just as Brazil's exports to us serve to pay for her imports from Great Britain.

Other types of bills of exchange. In our examples of bills of exchange we have mentioned only one type, the demand draft, although we have given two forms, the commercial demand draft, drawn by one commercial house on another, and the bankers' sight draft, drawn by one banker on another. We need not attempt a classification of the varied forms which the bill of exchange in foreign trade may take, but shall simply note the fact that such forms exist and are in common usage wherever the foreign bill of exchange is used. Attention should however be called to a very common form, the time or long bill of exchange, which requires payment a certain number of days after date, or after presentment, or *sight,* as it is called. If Richards & Green had sold its cotton to be paid for on the basis of a time bill of exchange, the bill might have read as follows:

No. 726 New York, Dec. 29, 1924
Sixty days after sight pay to the order of
The New York Bank
........One Thousand Pounds Sterling.......
(signed) Richards & Green
To Harris & Company, Ltd.
London, Eng.

The New York Bank will be unwilling to pay the same rate for this bill of exchange that it paid for the sight draft, because the bill will not be payable until sixty days after it has been presented to Harris & Company, Ltd., for acceptance. To determine its value at the moment of acceptance one must discount it for sixty days at the rate of discount prevailing in the London discount market. The rate which the New York Bank pays will take this fact into account in a very precise manner and will be lower than the sight rate, say $4.81 per pound sterling. When this bank sends the bill to the London Bank, it gives instructions to present it to Harris & Company, Ltd., for *acceptance*, which means that Harris & Company, Ltd., will write on the face of it " accepted and payable at the London Bank on such and such a date " and sign it, thereby formally acknowledging its obligation. The New York Bank may then request the London Bank to have the bill discounted and the proceeds placed to its credit, so that it may draw its own drafts against the credit, or it may conceivably desire to have it held until maturity, thereby earning the interest for the two months. Probably the largest portion of all our exports are paid for by time bills of exchange of one sort or another. They are therefore of considerable importance, but it should be remembered that they differ from the sight draft only in the time element, and that they introduce no new problems in the theory of foreign exchange.

The exchange market. The exchange rate responds in so sensitive a manner to the play of the forces of demand and supply and there is such a vast quantity of transactions in foreign exchange in New York and other centres, that it is frequently assumed that there must be an organized market similar to the stock and produce exchanges where the rate of exchange is determined. This is an error; there is in New York no formal organized exchange for dealing in foreign exchange. The bulk of the business is concentrated in the hands of a relatively few large banks and exchange houses. Each bank fixes the rates at which it will buy and sell exchange on all the financial centres with which it has dealings, but each house keeps so closely in touch with the activities of the other houses that there is seldom any material discrepancy between the rates charged by the various dealers on the same class of bills.

There is no official rate of exchange. Each bank posts the rates at which it is willing to buy and sell bills of exchange, but it frequently departs from its posted rates both on bills of large amounts and unquestioned character and on bills of dubious standing. In fact some bills will not be bought at all, but accepted for collection only. Since 1921 the Federal Reserve Bank of New York has been required to issue a daily report of the buying rate in New York for cable transfers on the important financial centres, and this possibly is the nearest approach to an official quotation. But it should be borne in mind that this rate is not a rate at which the Federal Reserve Bank itself is offering to buy cable transfers, but is one obtained from the regular dealers in exchange.

Fiat Money. When a monetary system is not on the gold standard, a new complication is introduced into the problem of foreign exchange. The mint par of exchange is no longer effective, and in its place must be put the ratio of the actual values of the two units, whatever they are. If as is usual the rates of exchange are still expressed as the price in one country of money payable in the other country, there is now no limit to the rise or fall of exchange rates. For example during and after the Civil War, when the United States was on a fiat money standard, our greenback dollar was worth less than a gold dollar. Since England was still on the gold standard, her pound was worth a good deal more than $4.86 in our paper money. The rate of exchange on London consequently rose far above the normal limit, or upper gold point. For years it stood above $6.50, and at one time it rose as high as $14.60.

After the World War the situation was reversed. The British pound was until 1925 fiat money and worth less than its former gold value, while the United States was on the gold standard. It did not require anything like $4.86 of our money to purchase a bill for one pound payable in London. At one time the rate was as low as $3.30. As late as January, 1923, it was $4.66.

In fact the whole foreign exchange situation was thrown by the monetary disturbances of the World War into confusion, from which it has even now only partially recovered.

EXERCISES

1. Mr. Jones, an exporter of grain from the United States to France, sells a shipment worth $10,000 to M. Guerne, agreeing to receive payment by drawing on the buyer.
    (*a*) Draw the bill of exchange, and trace out the steps by which the transaction is handled.
    (*b*) At par, what would be the face of the bill? 740

2. There are 20 shillings to the pound, and 12 pence to the shilling. From the information given in the text determine the mint par of exchange of a dollar in pence.

3. The gold franc contains 4.4803 grains of fine gold. Determine the mint par of exchanges:
    (*a*) in terms of cents per franc.
    (*b*) in terms of francs per dollar.
    (*c*) in terms of francs per pound sterling.

4. There are 480 grains to the ounce. The silver dollar of Hongkong contains 377.18098 grains of fine silver. The price of fine silver per ounce on December 31, 1924, was 66½ cents. Determine the price in American dollars of the silver dollar of Hongkong.

5. The result of the above gives the par of exchange between the American and the Hongkong dollar on December 31, 1924. Find from your newspaper the present quotation on silver bullion and determine the present par of exchange. What does this indicate with regard to the par of exchange between a country on a gold basis and one on a silver basis?

6. What is the present rate of exchange for demand drafts on Hongkong? What does this indicate as to the relation of demand and supply for exchange in Hongkong? as to the exports and imports to or from Hongkong?

7. Explain how our exports to Europe are used to make payment for our imports from the Far East.

CHAPTER XXVIII

THE INTERNATIONAL BALANCE OF PAYMENTS

Mercantile fallacies. Thus far we have not faced the question whether an excess of merchandise exports over imports always causes gold to flow into a country. Popular belief is firmly fixed on this point and has been for centuries. The Mercantilists were convinced that the prosperity of a country depended in large measure on the ability to pile up precious metals within the country, and they directed their domestic and foreign policy with this end in view. To them, as to their present-day counterparts, it seemed an obvious fact that if the exports of merchandise were greater than the imports of merchandise the balance must be paid in precious metals. The logical policy, it seemed, was to stimulate exports and repress imports in order that this desirable end might be attained. A " favorable " balance of trade, so called — an excess of merchandise exports over imports — was and still is considered by many a sign of national prosperity; an " unfavorable " balance of trade, a token that something ought to be done. There are three fallacies in the concept of a favorable balance of trade. The first is in conceiving the exports and imports of a country to be the only items in the balance sheet, the second lies in the concept of gold as the most desirable of all imports, and the third is the supposition that a steady inflow of gold can be produced by artificial means. The discussion which follows will revolve about these three points.

In this chapter we shall assume that a free gold market prevails in all of the countries concerned; *i.e.*, that each country's currency is either gold or redeemable in gold and that there are no legal obstacles in the way of the export and import of gold.

The balance of payments. If we would discover whether gold will be imported or exported by the bankers of the United States, we must have a record of all the transactions between individuals, corporations, and governments in this country and individuals, corporations, and governments in all other countries which have an influence upon the rate of exchange. Frequently in speaking of such transactions we say that America exports goods to England or that Brazil owes England a certain sum, but it should be kept in mind that reference is not to transactions between the respective governments, but to all transactions private and public between parties in the two nations.

In this record of international transactions we naturally include the exports and imports of merchandise, and, as has already been pointed out, we use a special term — the balance of trade — to denote the relationship of the exports to the imports of merchandise. But there are many items other than merchandise which must enter into the balance sheet of any nation and help to determine whether or not it will export or import gold. Every year the people of the United States incur indebtedness to the people of other countries because of a variety of services performed by them for us, and the same holds true of indebtedness of other countries with respect to the United States. The ships of Great Britain carry a large share of our foreign commerce, the greater proportion of the marine insurance premiums must be paid to insurance companies of Great Britain, foreign banks charge fees for handling the collection of bills of exchange in foreign markets. Each year we spend millions of dollars for securities issued by foreign governments and private concerns, and foreigners likewise invest in our securities. Payments must be made for these purchases and sales and also for the interest payments to the American holders of foreign securities and the foreign holders of American securities. American charitable organizations have been feeding and caring for the destitute in many foreign countries; our tourists overrun all the countries of Europe in the summer time, and in one way or another funds for their support must be placed abroad. Many other similar items might be mentioned which involve payment for services rendered or received.

Each one of such transactions will have some influence in determining the flow of gold, for each gives rise either to a demand or to a supply of bills of exchange and therefore helps determine the rate of exchange. Let us suppose for example that a corporation in the United States has interest payments to make to bondholders in England. It will designate some bank in England to serve as its agent in making payments to individual bond holders and will place this bank in possession of sufficient funds to enable it to redeem all the interest coupons presented for payment. In order to remit the necessary funds, the corporation will apply to some New York banker for a draft on London which it can remit to its London agent. In other words the obligation of the American corporation to make interest payments on London has given rise to a demand for bills of exchange in London and will exert the same influence on the rate of exchange as an importation of goods from England. In precisely the same way any transaction which places any person in England in the debt of an American will give rise to a supply of bills of exchange on London. If an American underwriting syndicate were to sell a British investment house a block of bonds in a new American corporation, it would thereby acquire the right to draw bills of exchange on the investment house; it would have bills to sell to American bankers and would so increase the supply of bills in the American market.

In order to have a concrete example let us assume that A and B are countries trading with each other exclusively and that the people of A in a particular year export $100,000,000 of goods to the people of B and import goods valued at $75,000,000. We find also that people in B have rendered for those of A shipping services worth $15,000,000, that the bondholders in B are owed $25,000,000 in interest, and that emigrants from B have made remittances of $5,000,000. On the other hand the people of A have performed services in carrying freight worth $5,000,000, they are owed $5,000,000 in interest on the foreign bonds which they hold, and emigrants from A have remitted $10,000,000. The balance sheet of A will be as follows:

INTERNATIONAL BALANCE SHEET. COUNTRY A

Credits	
Exports of merchandise	$100,000,000
Ocean freight	5,000,000
Interest	5,000,000
Emigrant remittances	10,000,000
	$120,000,000

Debits	
Imports of merchandise	$ 75,000,000
Ocean freight	15,000,000
Interest	25,000,000
Immigrant remittances	5,000,000
	$120,000,000

The bankers in A can secure credits of $120,000,000 in B by the purchase of the bills of exchange offered for sale by the exporters of merchandise, the carriers of freight, and the others who have the right to draw bills on foreigners; and the bankers can also sell $120,000,000 of bills of exchange to the importers, the corporations, the shippers, and the immigrants who have foreign payments to make. With this equilibrium of debits and credits, there is no occasion for the rate of exchange to be either much above or much below par; and since gold will not be imported until the rate falls to the gold importing point or exported till the rate rises to the gold exporting point, there is nothing to occasion a flow of gold either way.

Now let it be supposed, the other items in the example being the same, that the imports of merchandise into A had amounted to $90,000,000. In this case gold to the extent of $15,000,000 would have been exported, giving an item of this amount on the credit side and so balancing the account.

It is this record of all the debits and credits except gold in the international balance which we call the *balance of payments*, a term which thus includes all international transactions for which payment must be made, except gold imports and exports, which are the balancing items.

It will now be evident that it is the balance of payments, and not the balance of trade, which determines the flow of gold. In our illustrative balance sheet, country A had clearly a so-called favorable balance of trade; yet there was no import of gold, for the credits created by the excess of exports over imports of merchandise were required to pay for various services rendered by persons in B. The balance of payments was in equilibrium. In our second supposition there was actually an export of gold, in spite of the fact that country A still had a favorable balance of trade. It is the balance of payments which controls. If the sum of the debts which we owe to foreigners within a given period exceeds the sum of the debts of foreigners to us, gold will be shipped out of the country in settlement of the balance; and if foreigners owe us more than we owe them, gold will be shipped into the country.

Limits to the flow of gold. Returning to our example of countries A and B dealing exclusively with each other, let us suppose that A in a given year imports $100,000,000 in gold because of an extraordinarily large favorable balance of payments. The gold comes into the hands of the bankers in the first place and then is gradually put into circulation or else serves as a reserve for other forms of currency. Sooner or later its influence will be felt on prices, causing a rise in the general price level; and similarly the withdrawal of $100,000,000 from circulation in B will cause a decline in the general price level of that country. B therefore becomes a better market than A in which to purchase, and the people of A will incline to use more of the commodities and services of B in preference to their own country's more costly goods. Conversely the people of B will reduce their purchases of the high priced goods of A. The double result will be an increase in the imports and a decrease in the exports of A.

While gold was flowing into A the rate of exchange must have been down to the gold importing point. Now however there has come a transformation in the exchange market; the increase in imports means an increase in the demand for bills of exchange, while a decrease in the supply is caused by the falling off of exports. The simultaneous increase in demand and decrease in supply will have the effect of raising the rate of exchange to some point above the gold importing point in A, thus putting a stop to the further inflow of gold.

Likewise an outflow of gold from country A to country B would be only temporary, for the export of gold would reduce the price level in the gold exporting country and raise the price level of the gold importing country. In the first country the export of goods and services would be stimulated and the imports checked until the increase in the supply and the decrease in the demand for bills of exchange would cause the rate to fall below the gold exporting point, so stopping the outflow.

We conclude that gold cannot flow in one direction, either into or out of a country, in large amounts for an indefinite period, because the flow of gold provides its own automatic check through its effect on the price level, the exports and imports, and the rate of exchange. Temporary discrepancies in the balance of payments are settled from time to time in gold, but economic forces work in the direction of reducing the flow of gold to the minimum, and while in a given year a country may have considerable net imports or net exports of gold, consideration of a longer period will show the imports and exports of gold tending approximately to offset each other. The tendency over a period of time is for the debits and credits — the amounts owed by residents of a country to foreigners and the amounts owed by foreigners to them — to equal each other. We pay foreigners for goods and services by the export of our own goods and the rendering of services to foreigners; in the last analysis the money element becomes trifling, and international trade in the broadest sense reduces itself to barter.

A gold producing country. An exception to the general conclusions reached in the previous section appears in the case of a country which is normally devoting a considerable share of its energies to the production of gold. If A were suddenly to discover rich deposits of gold and were to start mining on a large scale, the first effect would be a rise in the price level. The higher prices in A as compared with B would, following the principles with which we are now familiar, lead to an increase of imports and a decrease in exports, and the exchange rate in A would rise to the gold exporting point and tend to remain there. Gold would normally and regularly be exported in payment for goods imported from B. Similarly B would year after year have large imports of gold. Gold, while still the balancing item in the commerce of the two nations, should be considered in the light of a domestic product. A happens to have advantage in the production of gold which leads her to devote her energies to the exploitation of the gold fields rather than, let us say, to the growing of wheat, and she finds it more profitable to secure her imports by the exportation of gold than by the exportation of wheat or some other commodity in the production of which her comparative advantage is less.

Recent gold importations of the United States. The reader is already aware of the fact that during the World War and the period subsequent to it the United States received several billion dollars in gold. That this was an abnormal situation can be seen from the fact that for the forty years prior to the war the net excess of imports over exports of gold amounted to only a little over $60,000,000. The war found the United States one of the few nations in a position to supply the Allies with the munitions, clothing, and food needed for the military and civil population. The enormous expansion in the demand for American export goods was not accompanied by a corresponding increase in the American demand for European goods, and the huge supply of bills of exchange on Europe soon drove the rate to the point where the importation of gold was indicated. The various governments tried to protect their stocks of gold by prohibiting its exportation but soon found themselves in the position where they had to export gold or do without the munitions and foodstuffs, and their needs were so imperative that the gold was surrendered. Ultimately, before their entire stock of gold had been absorbed by America, the European nations effected credit arrangements by which they might continue to receive from us goods far in excess of their ability to pay either in goods or in gold.

This however was an extraordinary situation and does not constitute an exception to the general principle that the flow of gold provides its own automatic check. Normally a rise in price reduces the quantity of goods which can be exported, but in a life and death struggle the purchase of certain goods is so inelastic that they will be taken whatever the price. Hence, although it is true that the vast importations of gold into the United States did in fact cause a rise in the price level which was without precedent for a country actually on a gold basis, it did not succeed in checking exports and causing a return flow of gold, because the prices of these goods rose in the European countries owing to the abnormal demand just as fast as they rose in America on account of the increase in gold. A new equilibrium was determined in which America held far more than her normal share of the world's gold supply.

Reaction of exchange rates upon imports and exports. Aside from the regulating force of the international flow of gold, the exchange rates themselves react upon imports and exports and tend to bring about an equilibrium. This relation will become clear if, placing ourselves in the position of the merchant whose business brings him into the exchange market as a buyer or seller of bills, we consider the effect of the exchange rates upon his transactions. Let us assume that this typical merchant is exporting to England and that sterling bills are selling at a high premium in New York. With exchange in this position, the merchant may make an extra profit from his sales in England. His normal trading profit is computed in the sale price of his goods and is therefore included in the face of the sterling bills which he draws against the English buyers. In turning his bills into dollars at the prevailing high rate of exchange, a second profit is derived from the fact that each pound sterling is worth more dollars than normally. Now if this merchant is in active competition with other exporters, as is the normal case, the profitable market for American goods in England will result in lower sale prices to the benefit of English buyers, and this situation will increase exportation, since it is axiomatic that more can be sold at a low than at a high price.

Obviously the position of the American importer of English goods is the reverse of this when sterling exchange sells at a premium in New York. Since payment for goods imported from England is made by the purchase of sterling bills in New York for remittance to the English exporter, the American importer finds his profit cut down by the premium on exchange. The premium becomes an additional cost in the business of importing English goods, and the prices of these goods will rise and their import diminish.

These illustrations should suffice to make it clear that the relation of the rates of exchange to the balance of payments is one both of effect and of cause. The rise and fall of the rates is controlled fundamentally by the balance of payments. But when an excess of imports forces foreign exchange to a premium in our market, the high rate tends to restrict our imports and encourage our exports and thus works toward a restoration of equilibrium in our foreign trade. Conversely when a large volume of exports makes the rates on foreign bills fall to a discount, our imports are stimulated and our exports restricted, which again tends to bring our trade into balance. In normal times the rates, confined as they are within the bounds of the gold points, can rise or fall so little that their influence upon the balance of payments may be imperceptible. It must be borne in mind however that this narrow range exists only because gold is permitted to flow in sufficient volume to offset the balance of payments; when gold is prevented from supplying this corrective, the rates of exchange will exert the influence needed to maintain equilibrium of exports and imports.

The effects of borrowing and lending. There are times when unusual transactions upset the equilibrium which has been established in foreign transactions. If either the people or the government of A were to borrow $100,000,000 through the sale of bonds to lenders in B, A would have a credit of that amount either with B's national treasury or with bankers in B. When a government borrows or when business men borrow from foreign lenders, what is wanted is generally equipment of one sort or another for use in the home country. Thus in an earlier period when American capitalists were selling large quantities of the stocks and bonds of American railway corporations to investors in England, the money was needed and used for the purchase of railway equipment in England. In such case the two items exactly balance each other. The bonds are placed on the credit side and the merchandise on the debit side of the balance sheet of the borrowing country, and of course on the opposite sides of the account for the lending country. The loan gives the borrowers a credit with the foreign bankers, upon which the borrowers draw bills of exchange to pay for the imported equipment.

The borrowers in A might however desire to use the proceeds of the loan in their own country, and with that end in view draw bills of exchange against these foreign credits and offer them for sale to the bankers in A. Such an increase in the supply of bills of exchange on B would quickly drive the exchange rate to the gold importing point, and the net result would be an importation of approximately $100,000,000. Again we have two balancing items — bonds against gold. But in this case sooner or later the imports of A from B would have to increase and its exports decrease, because of the influence of the imported gold on the respective price levels, as explained above. B may be a very wealthy country and quite able to make annual loans of several hundred millions of dollars, but it can only do this by exporting its goods or services. No country can have enough gold to permit of such a continual drain. We can therefore confidently expect in this second case to find B's exports increasing and her imports decreasing until the point is reached where enough of the exported gold returns to her to equalize the price levels again. In general it may be said with entire truth that a country makes loans with its goods and services and not with its gold, even though in the first instance gold may have been exported because of the loan.

The international balance sheet of the United States. Thus far we have been dealing principally with relations between hypothetical countries, a method of approach which, by simplifying the conditions of the problem, has the advantage of bringing out more clearly the fundamental principles sought. But that these relations are not inconsistent with the facts can easily be seen by a glance at the table below, which gives the figures compiled by the United States Department of Commerce, showing the balance of payments for the United States in 1924.

Balance of International Payments of the United States
1924
(millions of dollars)

Credits	
Exports of merchandise*	$4,621
Exports of silver	110
Ocean freights	76
Services to tourists	100
Interest and dividends	614
Sale of outstanding securities	319
Foreign bonds paid off	45
Principal of inter-allied debt	23
Increase on foreigners' bank balances	216
Exports of gold	62
Total credits	$6,186
Debits	
Imports of merchandise**	$3,651
Imports of silver	74
Ocean freights	68
Government payments abroad	5
Services to tourists	600
Charitable and missionary expenses	55
Immigrants' remittances	300
Interest and dividends	150
New foreign loans	795
Purchase of outstanding securities	114
Imports of U. S. paper currency	50
Imports of gold	320
Total debits***	$6,182

* This amount includes 30 million dollars for unrecorded parcel post packages.

** This amount includes 40 million dollars for smuggled liquors.

*** It is apparent that the credits and debits do not exactly balance, but it should be realized that many of the figures represent estimates only, and necessarily so, since there is no law compelling tourists to inform the government of the sums they spend abroad or compelling American shippers to certify as to the amounts they pay foreigners for ocean freight. It must also be remembered that there is no settlement date upon which a country's account for international payments must be balanced. The balance of payments is a continuing account, and even if our statistics were exact to the last dollar we should not expect an exact balance.

It will be noted that in 1924 we had a favorable balance of trade, since the exports of merchandise exceeded the imports of merchandise by $970,000,000. This is an instructive fact, because there has been so much public discussion of the inevitability of an unfavorable balance of trade in the near future as to lead to the erroneous conclusion that we now have such a balance. In 1924 we had also a favorable balance of payments, evidenced by a net import of gold and United States paper money (the balancing item) of $308,000,000. Normally the export or import of United States paper money would be so small in volume as to be a negligible factor, but the depreciation of foreign currencies and the restrictions placed on the movement of gold from foreign countries have made our paper money even more desirable than gold in such a country as Germany. We therefore have to include it with gold in the balancing item of the account.

The balance of trade once more. In view of the persistence of Mercantilist notions respecting the balance of trade and the real dread of the period in which an unfavorable balance of trade seems inevitable for the United States, it will be worth our while to summarize the previous discussion briefly and to offer some further explanations which may show that an unfavorable balance of trade is not the calamity which it is popularly believed to be.

A favorable balance of trade does not mean that gold will flow in. We have already discussed this point in sufficient detail to make clear that the only thing which can induce an inflow of gold is a favorable balance of payments. Normally gold will flow in and out only in very small quantities, simply to balance temporary indebtedness one way or the other. In a fifty year average, unless the situation is complicated by gold production or by disarrangement of the currencies of the various nations or by some other factor of like nature, the imports and exports of gold for any given country will be nearly equal. The tendency is for the debits and credits to offset each other without the importation or exportation of gold.

Is gold a peculiarly desirable import? Many persons would immediately answer this question in the affirmative. They experience a feeling of satisfaction when the statistics show that gold is coming in, they " view with alarm " an outflow of gold, and they would favor governmental action to encourage the import of gold.

In the first place it should be evident that gold obtained through international trade adds nothing to the total wealth of the nation, since it is obtained only by the sacrifice of other goods presumably of equal value. Even if the gold were obtained by an international loan, a debt is created equal to the value of the gold obtained. Those who believe gold imports are peculiarly favorable must base their belief upon the notion that gold is in some way superior to other forms of wealth or property. Is there any basis for this notion? Gold has certain important industrial uses, as in dentistry and the making of watches and jewelry; its principal other use is to serve as a medium of exchange. So far as its industrial uses are concerned, it will scarcely be maintained that gold is more essential to a people than a host of other commodities, wheat, meat, coal, iron, steel, lumber, and so on indefinitely. If a country lets these other things go in order to get gold, it presumably gains in utility through the import of gold. But it just as truly gains when it exports gold in order to obtain other things which it wants more. Gold differs not at all from other commodities in this respect, and the student who has any knowledge of economic principles need pursue this part of the inquiry no further. Certainly nothing could be more absurd than, by governmental action, to compel a people to give up things they want more in exchange for gold they want less.

As regards the use of gold as money, the important fact is that, within limits, the absolute quantity of money is pretty much a matter of indifference. Through the adjustment of prices, a given volume of exchanges may be consummated with either a small or a large quantity of money; what occurs is that the price level is low in the one case, high in the other case. When gold flows in and out naturally, in response to the economic forces which we have been studying, the quantity of currency and the price level in any given country adjust themselves to the international equilibrium, which is the condition most favorable to the people of the country in question. To acquire, by artificial means, a greater quantity of gold would raise the local price level, disturb the relations between debtors and creditors, disorganize the international balance of payments and the exchange market, and finally, unless prevented by other measures equally artificial, lead to the expulsion of the gold itself, all to no purpose.

Note that it is the whole currency system which comes into equilibrium with the system of other countries. Now it may well be that a given country needs more gold in its currency system, and that an inflow of gold for that purpose would be an advantage. But this is a problem of the monetary system, not of the country's foreign trade. Revise the monetary system so as to require a greater proportion of gold, and the gold will come in automatically through the channels of trade. Without such revision, the gold will not come or, if brought in, will not stay.

An unsound or disordered monetary system which requires the government to support a large credit currency may give to the treasury officials an anxious interest, such as the bankers always have, in the flow of gold. The resumption of specie payments after the paper money inflation of the Civil War was rendered easier by a turn in the balance of payments in our favor. An unusually large wheat crop coupled with disaster to the foreign crops came in time to check the outward flow of gold and start it flowing the other way. For twenty years thereafter the maintenance of the gold standard, in the face of inflation by greenbacks, silver dollars, and Treasury Notes of 1890, was a more or less heavy burden upon the Treasury and gave to the international movement of gold an interest which, except for the unsound monetary system, it need not have had. This is no denial of or exception to the conclusions at which we had arrived.

In general the flow of gold is an automatic thing, controlled by economic forces which result in a state of equilibrium which is for each nation the most favorable situation possible in view of all its other circumstances. There is nothing to set gold apart as a peculiarly desirable import, and nothing to be gained by forcing its import in excess of what would enter under the normal operation of international trade.

Is an adverse balance of trade a sign of national decline? We have seen that a favorable balance of trade does not necessarily mean an inflow of gold, that an unfavorable balance of trade does not necessarily imply an outflow, and further that even if it were possible to bring about the importation of gold by artificial means, the final result could only be a rise in the price level and a decline in the purchasing power of the gold unit. It may however still seem to some that a country which has an unfavorable balance of trade is losing its vigor and is on the road to economic ruin.

Yet England, Germany, and most of the highly industrialized countries have had unfavorable balances of trade for many years. Year after year these countries have imported a greater volume of commodities than they have exported. The United States on the other hand has long had a favorable balance of trade. Each year a portion of our exports have gone to pay for the use of foreign capital, for shipping and banking, and other services rendered to us. A decrease in our favorable balance of trade indicates that our dependence on those services is becoming less, and an unfavorable balance would mean that foreign countries were becoming dependent upon us, were paying for these services with their manufactured goods or raw materials. An unfavorable balance of trade is far from being a stigma of decadence; it simply indicates that the economic relations are different from those of a country with a favorable balance of trade. In particular it is apt to be the mark of an old established nation with a large accumulation of wealth.

EXERCISES

1. Construct a balance sheet from the following items, and determine the balance of trade and the balance of payments. Exports of merchandise $200,000,000; tourists' expenditures abroad, $150,000,000; shipping services rendered by this country, $150,000,000; purchase of foreign bonds, $50,000,000; emigrant remittances, $25,000,000; interest due to foreign stockholders, $10,000,000; imports of merchandise, $140,000,000; gold $——————?

2. The government of France is considering the question of refunding the debt to the U. S. government. If this debt should be refunded what effect would it have on the international balance? on the rate of exchange? Why?

3. If France were at the present time to repudiate her debt to the United

States, what effect would this have on the rate of exchange? the exports and imports? Why?

4. What effect does the annual payment of interest by the British government to the U. S. Treasury have on the sterling rate of exchange? on our exports and imports?

5. Suppose that the British government should decide to cease to pay these interest charges. Would that have any effect on the rate of exchange? on our exports and imports?

Suggestions for Further Reading on Part III. Money, Banking, and Exchange

I. MONEY AND BANKING

MOULTON, H. G. *The Financial Organization of Society.* Chicago, 1921
ROBERTSON, D. H. *Money.* New York, 1922
FOSTER, W. T. and CATCHINGS, W. *Money.* Boston and New York, 1923
JOHNSON, J. F. *Money and Currency.* Boston, 1905
NICHOLSON, J. S. *A Treatise on Money and Essays on Monetary Problems.* Fifth edition. London, 1901
HEPBURN, A. B. *History of Currency in the United States.* New York, 1924
WATSON, D. K. *History of American Coinage.* New York, 1899
LAUGHLIN, J. L. *History of Bimetallism in the United States.* New York, 1900
MITCHELL, W. C. *History of the Greenbacks.* Chicago, 1903
WHITE, H. *Money and Banking.* Fifth edition. Boston, 1911
HOLDSWORTH, J. T. *Money and Banking.* New York, 1914
MOULTON, H. G. *Principles of Money and Banking.* Chicago, 1917
PHILLIPS, C. A. *Readings in Money and Banking.* New York, 1916
WESTERFIELD, R. B. *Banking Principles and Practice.* New York, 1924
DUNBAR, C. F. *Chapters on the Theory and History of Banking.* Fourth edition. New York, 1922
PHILLIPS, C. A. *Bank Credit.* New York, 1920
CONANT, C. A. *History of Modern Banks of Issue.* Fourth edition. New York, 1909
CANNON, J. G. *Clearing Houses.* New York, 1900
ROBB, T. B. *The Guaranty of Bank Deposits.* Boston, 1921
ECKHARDT, H. M. P. *A Rational Banking System (Branch Banking).* New York, 1911
REED, H. *The Development of the Federal Reserve Policy.* Boston and New York, 1922
DEWEY, D. R. *Financial History of the United States.* Eighth edition. New York, 1922 (See Table of Contents)
NOYES, A. D. *Forty Years of American Finance.* New York, 1909

II. Money and Prices. The Value of Money

Fisher, I. *The Purchasing Power of Money.* New York, 1911

Layton, W. T. *An Introduction to the Study of Prices.* London, 1914

Mitchell, W. C. *Gold, Prices, and Wages under the Greenback Standard.* Berkeley, 1908

Fisher, I. *The Making of Index Numbers.* Boston, 1922

Mitchell, W. C. *The Making and Using of Index Numbers.* Bulletin of the United States Bureau of Labor Statistics, No. 284. October, 1921

Fisher, I. *Stabilizing the Dollar.* New York, 1920

III. The Business Cycle

Adams, A. B. *Economics of Business Cycles.* New York, 1925

Mitchell, W. C. *Business Cycles.* Berkeley, 1913

Moore, H. L. *Economic Cycles: their Law and Causes.* New York, 1914

——— *Generating Economic Cycles.* New York, 1923

Edie, L. D. (Editor) *The Stabilization of Business,* by Wesley C. Mitchell and others. New York, 1924

Business Cycles and Unemployment. Report and Recommendations of a Committee of the President's Conference on Unemployment, including an investigation made under the auspices of the National Bureau of Economic Research. New York, 1923

Hastings, H. B. *Costs and Profits: their Relation to Business Cycles.* Boston, 1923

Jordan, D. F. *Business Forecasting.* New York, 1921

Vance, R. *Business and Investment Forecasting.* New York, 1922

IV. Interregional Trade

Bastable, C. F. *Theory of International Trade.* London, 1900

Brown, H. G. *Principles of Commerce.* New York, 1916

——— *International Trade and Exchange.* New York, 1914

Bishop, A. L. *Outline of American Foreign Commerce.* Boston, 1923

Griffin, C. E. *Principles of Foreign Trade.* New York, 1924

Wolfe, A. J. *Theory and Practice of International Commerce.* New York, 1919

Nourse, E. G. *American Agriculture and the European Market.* New York, 1924

Filsinger, E. B. *Exporting to Latin America.* New York, 1916

Eldridge, F. R. *Oriental Trade Methods.* New York, 1924

V. FOREIGN EXCHANGE AND COMMERCIAL RELATIONS

ESCHER, F. *Foreign Exchange Explained.* New York, 1917
FURNISS, E. S. *Foreign Exchange.* Boston, 1922
WHITAKER, A. C. *Foreign Exchange.* New York, 1919
EDWARD, G. W. *International Trade Finance.* New York, 1924
CROSS, I. B. *Domestic and Foreign Exchange.* New York, 1923
SPALDING, W. F. *The London Money Market.* London, 1922
BOGGS, T. H. *International Trade Balance in Theory and Practice.* New York, 1922
CULBERTSON, W. S. *International Economic Policies.* New York, 1925
LAY, T. H. *The Foreign Service of the United States.* New York, 1925

III Influence of environment.
1. Dark skin of tropic dwellers.
2x Lazyness " " "
3. White skin of Eskamos.
4. Ambition of "
5. Eyes of fish at mamouth cave.
6. Criminals from criminal surroundin
7xxxx False modesty. BLAU
8.

Geographical ~~geography~~ mythology

1 That compass points to north pole.

2 That H — is straight down

3 That sunshine on 2nd Feb mean 6 weeks winter.

4. China is straight through.

5. The nearer the sun the warmer the weather.

I Ways of Gaining Living

1. Economic.

   Farming & manufacturing

2 Neutral.

   Trading & gathering fruit.
   robbery

3. Uneconomic

   wrecking houses &

II Economic geography aids business ~~man~~ thus; Fur coats would have no sale in the tropics yet here there is a big demand for them. Loca has made a demand for warmer wrap

IV Instinct to save results from generations finding that supplies are needed to carry over the periods of drought.

V

Jerry

Monapoly { Legal
Natural
Capitalistic.

Monopoly means regulation of supply to determine Price.

9/

10. ~~Why would this affect the price~~

Text. from last. Chap X

Jan 11 '30 Japan on Gold Basis

22 countries have done same

since war)

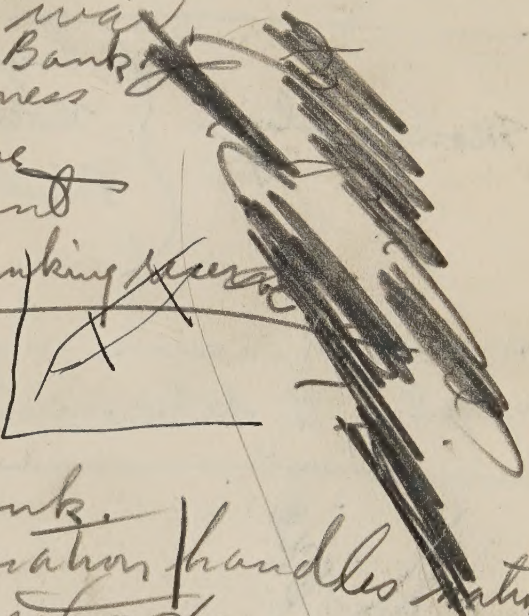
Central Bank

Bank business

1. note issue

2 rediscount

3. nation Banking reserve



Eng Bank.

like corporation handles nations money

no connection to government

all banks keep large reserve in 1st bank

note issue dept } Eng Bank.

Banking dept

notes backed by government obligations

or gold

Elasticity furnished by elastic deposit

& extraordinary flow of gold to London.

French bank. private but gov. by gov.

notes on rediscount, loans to gov. & people

makes representative money

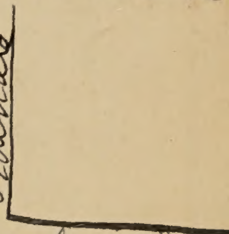
propo
liquidation
recovery.

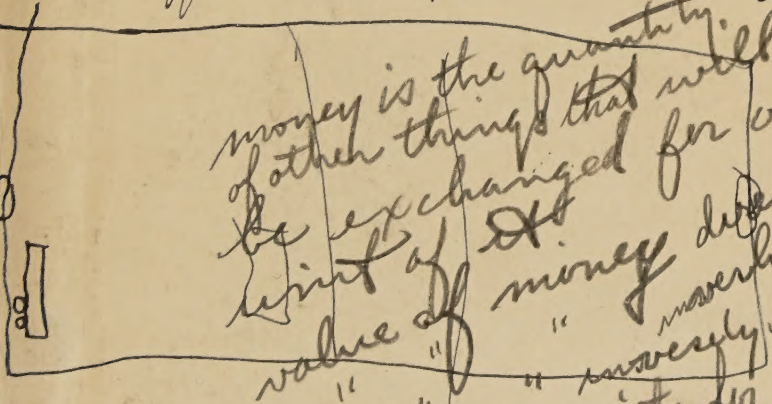
Determining charges on freight ~~on~~ depends on Cost of Service and Value of Service.

Candian bank. – no central bank notes on liens with priority to liabilities also safety fund of 5% of note issue

15
5
1.5
5
100

Supply of any good is a schedule of quantities of that good which sellers would offer at all prices.





money is the ~~quantity~~ of other things that will be exchanged for one unit of ~~it~~

value of money ~~directly with~~ inversely to quantity
" " " inversely with velocity
" " " opposite of level of prices

Law of Supply

In Given market at given time, the quantity of any good which would be offered for sale when not a fixed quantity varies directly with the price.